最优化原理

胡适耕　施保昌

华中理工大学出版社

图书在版编目(CIP)数据

最优化原理/胡适耕 施保昌
武汉:华中理工大学出版社,2000年5月
ISBN 7-5609-2193-0

Ⅰ. 最…
Ⅱ. ①胡… ②施…
Ⅲ. 最优化-高等学校-教材
Ⅳ. O224

最优化原理

胡适耕 施保昌

责任编辑:李立鹏 封面设计:刘 卉
责任校对:张 欣 责任监印:张正林

出版发行:华中理工大学出版社
武昌喻家山 邮编:430074 电话:(027)87545012

经销:新华书店湖北发行所

录排:华中理工大学出版社照排室
印刷:华中理工大学出版社印刷厂

开本:850×1168 1/32 印张:8.375 插页:2 字数:176 000
版次:2000年5月第1版 印次:2000年5月第1次印刷 印数:1—2 000
ISBN 7-5609-2193-0/O·213 定价:11.80元

内容提要

本书对所谓无限维最优化理论的基本内容提供一个系统的处理．全书共8章．头两章概括了阅读本书主要内容所需的预备知识，其中包括基本的泛函分析结果与非光滑分析．随后各章阐述最优化理论的基本论题：不等式系统与择一定理、一阶与高阶最优性条件、对偶理论、向量最优化等．本书一方面以紧凑的形式概括了最优化理论的标准内容，同时介绍了较多的新近研究成果，其中包括作者本人的一些结果．这部分内容涉及近年来引起广泛关注的一些研究领域，因而可能为有研究兴趣的读者架设起从基础理论通向研究前沿的桥梁．对于数学系的高年级大学生及有关理工科专业的硕士生，本书略加删节之后可作为教材使用．在当代科学发展进程中，对于最优化理论的日益广泛与紧迫的需要，已成为一种引人注目的潮流；有这种需要的科技工作者，将发现本书可提供一些有用的理论工具．

Abstract

The purpose of this book is to provide a systematical treatment for the modern theory of so-called infinite-dimensional optimization. The book includes eight chapters. The first two chapters give some necessary preliminary knowledge for reading the book. The other chapters contain some basic topics of optimization: alternative theorems; first-order and higher-order optimality conditions; duality theory; vector optimization, etc. On the one hand the book contains a comprehensive treatment for standard results of optimization theory, on the other hand it provides some new concepts and many recent results which are the subject of great importance and much activity. The book can serve as textbook or reference for graduate students. It can also be consulted by relevant teachers, researchers, natural scientists and engineers.

写在“研究生用书”出版10周年

在今天，面对科技的迅速发展，知识经济的已见端倪，国际竞争也日趋激烈，显然，国家之间的竞争是国家综合实力的竞争，国家综合实力的竞争关键是经济实力的竞争，而经济实力的竞争关键又在于科技(特别是高科技)的竞争，科技(特别是高科技)的竞争归根结底是人才(特别是高层次人才)的竞争，而人才(特别是高层次人才)的竞争基础又在于教育。“百年大计，教育为本；国家兴亡，人才为基。”十六个字、四句话，确是极其深刻的论断。目前，国际形势清楚表明：我们国家的强大与民族的繁荣，主要立足于自己，以“自力更生”为主；把希望寄托于他人，只是一种不切实际的幻想。这里，我们决不是要再搞“闭关锁国”，搞“自我封闭”，因为那是没有出路的；我们强调的是要“自信，自尊，自立，自强”，要“自力更生”为主，走自己发展的道路。

显然，知识经济最关键的是人才，是高层次人才的培养，而作为高层次人才培养的研究生教育就在一个国家的方方面面的工作中，占有十分重要的战略地位。可以说，没有研究生教育，就没有威伟雄壮的科技局面，就没有国家的强大实力，就没有国家在国际上的位置，就会挨打，就会受压，就会被淘汰，还说什么知识经济与国家强大?!

“工欲善其事，必先利其器。”教学用书是教学的重要

基本工具与条件。这是所有从事教育的专家所熟知的事实。所以，正如许多专家所知，也正是原来的《“研究生用书”总序》中所指出，研究生教材建设是保证与提高研究生教学质量的重要环节，是一项具有战略性的基本建设。没有研究生的质量，就没有研究生教育的一切。

我校从1978年招收研究生以来，即着力从事于研究生教材与教学用书的建设。积十多年建设与实践的经验，我校从1989年起，正式分批出版“研究生用书”。第一任研究生院院长陈珽教授就为之写了《“研究生用书”总序》，表达了我校编写这套用书的指导思想与具体要求，“要力求‘研究生用书’具备科学性、系统性、先进性”。后三任研究生院长，也就是各任校长黄树槐教授、我本人和周济教授完全赞同这一指导思想与具体要求，从多方面对这套用书加以关心与支持。

我是十分支持出版“研究生用书”的。早在1988年我在为列入这套书中的第一本，即《机械工程测试·信息·信号分析》写“代序”时就提出：一个研究生应该博览群书，博采百家，思路开阔，有所创见。但这不等于他在一切方面均能如此，有所不为才能有所为。如果一个研究生的主要兴趣与工作不在“这一特定方面”，他也可以选择一本有关的书作为了解与学习这方面专业知识的参考；如果一个研究生的主要兴趣在“这一特定方面”，他更应选择一本有关的书作为主要学习用书，寻觅主要学习线索，并缘此展开，博览群书。这就是我赞成为研究生编写系列教学用书的原因。

目前，这套书自第一本于1990年问世以来，已经渡

过了10个春秋，出版了8批共49种，初步形成规模，逐渐为更多读者所认可。在已出版的书中，有15种分获国家级、部省级图书奖，有16种一再重印，久销不衰。采用此套书的一些兄弟院校教师纷纷来信，赞誉此书为研究生培养与学科建设作出了贡献，解决了他们的“燃眉之急”。我们感谢这些赞誉与鼓励，并将这些作为对我们的鞭策与鼓励，“衷心藏之，何日忘之?!”

现在，正是江南春天，“最是一年春好处”。华工园内，红梅怒放，迎春盛开，柳枝油绿，梧叶含苞，松柏青翠，樟桂换新，如同我们的国家正在迅猛发展、欣欣向荣一样，一派盎然生机。尽管春天还有乍寒时候，我们国家在前进中还有种种困难与险阻，来自国内与来自国外的阻挠与干扰，有的还很严峻；但是，潮流是不可阻挡的，春意会越来越浓，国家发展会越来越好。我们教师所编的、所著的、所编著的这套教学用书，也会在解决前进中的种种问题中继续发展。然而，我们十分明白，这套书尽管饱含了我们教师的辛勤的长期的教学与科研工作的劳动结晶，作为教学用书百花园中的一丛鲜花正在怒放，然而总会有这种或那种的不妥、错误与不足，我衷心希望在这美好的春日，广大的专家与读者，不吝拔冗相助，对这套教学用书提出批评建议，予以指教启迪，为这丛鲜花除害灭病，抗风防寒，以进一步提高质量，提高水平，更上一层楼，我们不胜感激。我们深知，“一个篱笆三个桩”，没有专家的指导与支持，没有读者的关心与帮助，也就没有这套教学用书的今天。我衷心祝愿在我们学校第三次大发展的今天，在百年之交与千年之交的时候，这套教学用书会以更

雄健的步伐，走向更美好的未来。

诗云："嘤其鸣矣，求其友声。"这是我们的心声。

中国科学院院士
华中理工大学学术委员会主任
杨叔子
于华工园内
1999 年 5 月 15 日

前　　言

近 20 年间，如果说有一个词，它既回响于科学殿堂，又流行于社会各界乃至市井街头；既闪耀着科学思想的火花，又融会着公众常识的直觉．那么，这个词就是**优化**或**最优化**(Optimization)．无论工程设计，生产经营，投资决策，经济运行，人才管理，社会结构，……，人们都在追求一种至上的境界，追求一种"极致"，这种普遍的冲动最终朝向最优化．而且，自然界本身早已按照最优化的原则决定其存在形态与演化方式．综观自然与社会，人们确信：最优化乃是任何事物趋于平衡时无可逃遁的一条规则．如此普遍的规则不可能没有与之相应的伟大数学理论！作为一门数学学科的"最优化"，非同寻常地热乎起来，乃为自然之势．

最优化问题可简单地表述为：在给定条件下求一函数的极值点．在这种意义上，最优化理论源远流长．然而，只是到 20 世纪下半叶，关于最优化的一些基本结论才被发现．因此，作为一门独立数学学科的最优化，乃是最近几十年间数学迅速发展的产物．

对一个最优化问题的解答自然分为两个部分．首先，必须回答该问题是否有解及其解集具有何性质．这方面的研究构成"最优化理论"，其基本内容正是本书所要介绍的．其次，对于一个确知其有解的最优化问题，具体求出其(准确或近似)解无疑有重大实际意义．一些愈来愈强有力的算法的涌现，正是近几十年来最优化方法的主要成就之一．这方面的内容将在本书的续著中加以介绍．

本书以 Banach 空间作为处理最优化问题的基本空间框架．这看来是一种较合理的选择，它既不像 Euclid 空间那样失之过窄，也不像拓扑向量空间那样失之过宽．因此，本书所处理的实际上是"无限维最优化"．本书尽了最大努力来证明：关于"有限维最优化"的许多结论，在某种更为简洁(因而也更自然)的形式下，也适用于无限维最优化．这是一个极令人鼓舞的事实，它不仅带来了具有高度概括性的统一理论，而且为最优化理论应用于数学物理及控制

理论等领域开辟了道路.

本书第一章给出全书所需的预备知识. 对于"无限维最优化"这一课题来说,泛函分析无疑是必需的基本工具. 第二章主要是为讨论"非光滑最优化"作准备的. 部分地应最优化理论之需要而发展起来的"非光滑分析",自身已成为一个内容丰富的独立学科,其应用价值已超出最优化理论之外. 第三章所处理的"择一定理"包含了一系列互有联系的定理,它们在最优化理论中起着十分独特的作用,被誉为"最优化理论的基石". 读者将会发现,该章是作者最着力的部分之一,大部分结果被推进到迄今所知的最强的形式,不少结果属于作者且是第一次发表. 我们认为,即使作为一个独立的研究课题,择一定理亦有非同寻常的引人之处;它所独具的优美逻辑形式在数学中是很典型的. 第四、五两章无疑是本书的中心内容,其中概括了"单目标最优化理论"的主要基本结果. 大部分结果基于近期文献,但作了尽可能的改进与整理,因而往往更具一般性或形式更为简洁. 作者相信有一部分结果实质上是新的,当然,其合理性与价值仍有待读者审验. 第六章所处理的向量最优化可分为两部分:其一是"标量最优化"的直接推广,这部分内容虽不可缺少,但不是最有意思的,它反映了一个理论在其发展进程中的例行扩张. 另一部分则是真正体现向量序特色的内容,它包含许多远未完全克服的重大困难,因而对于喜欢应付严重挑战的研究者更具吸引力. 对于这样一个问题与结果都在不断涌现的研究领域,本书作者只能谨慎地选择若干较成熟的材料,以作为对有兴趣读者的初步导引. 从逻辑上看,第七章是第四章的自然发展. 但"高阶问题"无疑具有更大的难度,因而不可期望与"一阶问题"同样成熟的结果. 实际上,已有多种处理高阶最优性条件的方法,本书只是选择了其中的一种而已. 本书所用的"高阶变分导数"不失为一个有效工具,以至作者对之情有独钟. 但作者无意贬斥其他可行的选择.

在任何意义上,本书前七章都不是最优化理论的一个全面总

结，因而自然有许多重要内容未能涉及．最后的一章似乎应扮演“补遗”的角色，但该章短短的三节很难起到这种作用．从本书所能容许的篇幅来说，作者至此已该搁笔了，况且，对于那些非专章不能论其详的重要专题（例如灵敏度分析），简略的描述无异于一种轻率．

由于已有预备性的第一章，本书基本上是自给自足的．作者并不追求各章的独立性，而是力图将全书组织成一个前后连贯的整体．我们认为，不同部分之间的密切联系有助于读者理解本书的内容．作者尽了很大努力去简化对内容的处理，其中包括采用一套特别有效的符号．这可能会影响本书的可读性，但要在一本区区二十余万字的书中容纳与本书相当的内容，任何作者在可读性与简洁性之间大概都很少有选择的余地．这使作者能够聊以自慰．凡想顺利阅读本书的读者，务必浏览一下本书卷首的“记号与约定”．

本书的大部分内容曾在华中理工大学数学系“非线性分析讨论班”上讨论过，作者衷心感谢讨论班成员对于撰写本书的热诚支持．著者也衷心感谢华中理工大学出版社领导及编辑为本书顺利出版而作的大量努力．

作者

1998.6.1

记号与约定

A° ($\overline{A}$,∂A)：集 A 的内部(闭包，边界).

$A^c=X\backslash A(A\subset X)$：$A$ 的补集.

A^*：集 A 的对偶锥或算子 A 的对偶算子.

$A^\perp$：集 A 的正交补.

$B_r(x)=B(x,r)=\{y: |y-x|<r\}$.

$\overline{B}_r(x)=\overline{B}(x,r)=\{y: |y-x|\leqslant r\}$.

C^r：r 次连续可微函数类.

C^{1-0}：局部 Lipschitz 函数类.

$C(A)$：A 上的实连续函数之全体.

co A：集 A 的凸包；$\overline{\text{co}}\ A=\overline{\text{co}\ A}$.

cone A：集 A 生成的锥；$\overline{\text{cone}}\ A=\overline{\text{cone}\ A}$.

cl $A=\overline{A}$.

$D(F)=D_F$：F 的定义域.

$D_f=\{x: f(x)<\infty\}$.

$\mathrm{D}f$：f 的 G-导数.

$\mathrm{D}^n f(x,h)=\frac{\mathrm{d}^n}{\mathrm{d}t^n}f(x+th)|_{t=0}$；$\mathrm{D}f(x,h)=\mathrm{D}^1 f(x,h)$.

$\mathrm{D}_\pm f(x,h)$：单侧方向导数.

$d(A,B)$：A,B 间的距离；$d_A(x)=d(x,A)$.

$\Delta f(x,h)=f(x+h)-f(x)$.

δ_A：集 A 的指示函数.

$\partial f(x)$：f 在 x 的次微分或 Clarke 次微分.

epi F：F 的上图.

$f^{(r)}$：r 阶 Fréchet 导数；$f'=f^{(1)}$.

f^*：f 的共轭函数.

$F_D(x)=F(D,x)$：集 D 在 x 的可行锥.

$F_f(x)$：f 在 x 的降锥.

Gr F：F 的图形.

$\text{int } A = A^\circ$.

$\inf A = \min \overline{A}$.

$J=[0,1]$(除少数指明的例外).

$K_D(x)=K(D,x)$：集 D 在 x 的 Bouligand 切锥.

LCS：局部凸拓扑向量空间.

Lip：Lipschitz 函数类.

Lip f：f 的 Lipschitz 模数.

$L(X,Y)$：从 X 到 Y 的连续线性算子之全体.

$L(x,\lambda,\mu)$或 $L(x,\rho,\lambda,\mu)$：Lagrange 函数.

lsc：下半连续.

$N(A)$：算子 A 的零空间.

$N_D(x)$：集 D 在 x 的法锥.

$R(F)$：F 的值域.

$\mathbf{R}_+=[0,\infty)$；$\mathbf{R}_-=(-\infty,0]$；$\overline{\mathbf{R}}=(-\infty,\infty]$.

$\mathbf{R}^n_+=(\mathbf{R}_+)^n$；$\mathbf{R}^n_-=-\mathbf{R}^n_+$；$\dot{\mathbf{R}}^n_+=\text{int } \mathbf{R}^n_+$.

$S(X)$：X 中的单位球面.

$S^n=S(\mathbf{R}^{n+1})$.

s_A：集 A 的支承函数.

span A：A 生成的子空间，$\overline{\text{span}}\ A=\overline{\text{span}\ \overline{A}}$.

$\sup A=\max \overline{A}$.

$T_D(x)$：集 D 在 x 的 Clarke 切锥.

usc：上半连续.

VOP：向量最优化问题.

wlsc：弱下半连续.

wmax A：A 的弱极大元之全体；wmin $A=-$wmax$(-A)$.

winf A = wmin $\overline{A}$；wsup A = wmax $\overline{A}$.

wusc：弱上半连续.

X,Y,X,W 等：实 Banach 空间，有时为 LCS.

X^*：X 的对偶空间.

2^X：X 的子集之全体.

$[x,y]=\{tx+(1-t)y:0\leqslant t\leqslant 1\}$.

$\langle u,x\rangle=u(x)(u\in X^*,x\in X)$.

几点说明

1° 括号的用法 方括号内记述意义平行(或对偶)的条件与结论；圆括号内是说明性插入语.

2° 引证 §1.1(1)表§1.1中式(1)，1.1.1(i)表定理(或命题)1.1.1的结论(i)；[1,Th.1]表文献[1]中定理1，余类推.

3° 指标的用法 指标(如 i,j,k,l)的变程(如 $1\leqslant i\leqslant n$)在一节中通常不变且只注明一次；出现于 $\sum,\prod,\bigcup,\bigcap$ 下的指标通常省略；给定 $x\in\mathbf{R}^n$，$f:\mathrm{D}\to\mathbf{R}^n$，自动认定 $x=(x_i)$；$f=(f_i)(1\leqslant i\leqslant n)$；对 $I\subset\{1,2,\cdots,n\}$，记 $f_I=\{f_i:i\in I\}$.

4° 集记号 若 $f(x,y)$，$F(x,y)$ 分别为 $A\times B$ 上的函数与多值函数，则记 $f(A,B)=\{f(x,y):x\in A,y\in B\}$，$F(A,B)=\bigcup_{x\in A,y\in B}F(x,y)$，如 $A+B=\{x+y:x\in A,y\in B\}$，$\mathbf{R}K=\{rx:r\in\mathbf{R},x\in K\}$，$\mathrm{D}f(x,A)=\{\mathrm{D}f(x,y):y\in A\}$，$\langle U,A\rangle=\{\langle u,x\rangle:u\in U,x\in A\}$.

5° 不等号的用法 对集 A,B，约定 $A\leqslant B\Leftrightarrow\forall a\in A,b\in B:a\leqslant b$；$A<B$ 与 $A\ll B$ 仿此. 对映射 F,G，约定 $F\leqslant G\Leftrightarrow\forall x\in D_F(=D_G):Fx\leqslant Gx$；$F<G$ 仿此.

6° 范数记号 任一 Banach 空间及其对偶空间中的向量范数与有界线性算子范数皆记为 $|\cdot|$，$|\cdot|_0$ 记为 sup 范数.

7° 锥记号 Y_+ 记 Banach 空间 Y 中一给定的闭凸锥，未加说

明时在 Y 中使用由 Y_+ 导入的序$\leqslant$;$Y_-=-Y_+$,$\mathring{Y}_+^*=\text{int }Y_+^*$;任给锥 K,约定 $K_-=-K$.

8° 极限记号 $\lim\limits_n x_n=\lim\limits_{n\to\infty} x_n$;$\lim\limits_{t\downarrow 0}\varphi(t)=\lim\limits_{0<t\to 0}\varphi(t)$;$\rightharpoonup$与$\overset{*}{\rightharpoonup}$分别记弱收敛与弱*收敛.

9° 函数记号 未注明时,字母 F,G,H 等表示:$F:X\to 2^W$,$G:X\to 2^Y$,$H:X\to 2^Z$.$f:X\to W$(或 $f:X\to \mathbf{R}$),$g:X\to Y$,$h:X\to Z$,$T=f'(\bar{x})$,$A=g'(\bar{x})$,$B=h'(\bar{x})$.$f|M$ 记函数 f 在集 M 上的限制.$\forall\rho\in W^*$,令 $f_\rho=\rho f$;g_λ,h_μ 仿此.

10° $\forall\, t\in[0,1]$:$t'=1-t$.

11° 零记号 数零、零向量、零函数、零映射、积空间(如 $X\times Y$)中的零元$(0,0)$等皆记作 0.

12° “$x\to x_0$ 时 $f(x)$有性质 P”意味着:当 x 充分接近于 x_0 时 $f(x)$有性质 P.例如,“$t\downarrow 0$ 时 $x_t\in A$”意指:当 $t>0$ 充分小时 $x_t\in A$.

13° * 记号 $g:X\to Y$ 为 * 拟凸(或 * lsc)$\Leftrightarrow\forall\,\lambda\in Y_+^*$:$\lambda g$ 为拟凸(或 lsc);其余依此类推.

目　录

第一章　预备知识

本章概述最优化理论所必需的基本分析工具，凡属标准的结果皆省去证明.

§1.1　基本泛函分析结果

本书中，X,Y,Z,W 等记取定的实 Banach 空间（简称 B-空间），其中的范数及对应的对偶空间范数、算子范数皆记作 $|\cdot|$. 任给 $A\subset X, U\subset X^*$，约定

$$\langle U,A\rangle=\{u(x):u\in U,\ x\in A\}.$$

$\mathrm{co}A$ 与 $\mathrm{span}A$ 分别记 A 的凸包与 A 生成的子空间；令 $\overline{\mathrm{co}}A=\overline{\mathrm{co}A}$，$\overline{\mathrm{span}}A=\overline{\mathrm{span}A}$. 约定 $A\leqslant B\Leftrightarrow\forall a\in A,\ b\in B:a\leqslant b$，只要 $a\leqslant b$ 恒有定义；$A<B$ 仿此. 认定 $(X\times Y)^*=X^*\times Y^*$，这意味着 $\forall\varphi\in(X\times Y)^*$，有唯一 $(u,v)\in X^*\times Y^*$，使得 $\varphi(x,y)=u(x)+v(y)(x\in X,y\in Y)$. 特别，$(X\times\mathbf{R})^*=X^*\times\mathbf{R}$. $\rightharpoonup$ 与 $\overset{*}{\rightharpoonup}$ 分别记弱收敛与弱 * 收敛.

1.1.1 Hahn-Banach 定理　设 A 是 X 的子空间.

(i) 若 u 是 A 上的线性泛函，p 是 X 上的次线性泛函（次线性意味着 $p(\alpha x)=\alpha p(x)$，$p(x+y)\leqslant p(x)+p(y)$，$\alpha\geqslant 0$，$x,y\in X$），$u\leqslant p|A$，则有 X 上的线性泛函 v，使得 $v|A=u$，$v\leqslant p$.

(ii) 若 u 是 A 上的连续线性泛函，则 $\exists v\in X^*:v|A=u$，且

$$|v|=|u|_A\overset{\mathrm{def}}{=\!=}\sup_{x\in A,|x|\leqslant 1}|u(x)|.$$

(iii) 若 $0\neq x\in X$，则 $\exists u\in X^*:u(x)=|x|$ 且 $|u|=1$.

由此推出：$x \rightharpoonup y \Leftrightarrow \forall u \in X^*: u(x)=u(y)$.

1.1.2 分离定理 设 $A,B\subset X$ 是非空凸集.

(i) 若 $A^\circ \neq \varnothing, A^\circ \cap B=\varnothing$，则 $\exists u\in X^*, r \in \mathbf{R}: u(A^\circ) < r \leqslant u(B), u(A) \leqslant u(B)$.

(ii) 若 $A \cap B = \varnothing$，A 是紧集，B 是闭集，或 A,B 皆为开集，则存在 $u \in X^*, r \in \mathbf{R}: u(A) < r < u(B)$.

(iii) 若 $\dim X < \infty, A \cap B = \varnothing$，则$\exists u \in X^*\backslash\{0\}: u(A) \leqslant u(B)$.

1.1.2 适用于 X 为 **LCS**（＝局部凸拓扑向量空间）的情况. 因 X^* 依弱*拓扑是 LCS 且其拓扑对偶就是装备弱拓扑的 X，故以 X^* 代 X 从 1.1.2 得出：

1.1.3 推论 设 $P,Q \subset X^*$ 是非空弱*闭凸集，P 弱*紧，$P \cap Q = \varnothing$，则 $\exists x \in X, r \in \mathbf{R}: \langle P,x\rangle < r < \langle Q,x\rangle$. 特别，若 $Q \subset X^*$ 是非空弱*闭凸集，$u \in X^*\backslash Q$，则 $\exists x \in X: u(x) < \langle Q, x\rangle$.

以上结果在本书中将反复用到.

1.1.4 Mazur 定理 设 $A \subset X$ 为凸集，$\overline{A}^w$ 记 A 的弱闭包，则 $\overline{A} = \overline{A}^w$；$A$ 是闭集 $\Leftrightarrow$ A 是弱闭集.

当 $\dim X = \infty$时，在某种意义上 X 中的弱闭集远少于闭集，因此判定“弱闭”更困难. 1.1.4 表明，对于凸集不存在这一问题. 若 X 是**自反空间**（即 $X = X^{**}$），则 1.1.4 亦可用于 X^* 中的凸集.

1.1.5 Banach-Alaoglu 定理 设 $U \subset X^*$，则 U 相对弱*紧 $\Leftrightarrow U$ 有界 $\Leftrightarrow U$ 等度连续. 后者意味着 $\forall \varepsilon > 0, \exists \delta > 0, \forall u \in U, \forall x \in B_\delta(0): |u(x)| < \varepsilon$. 特别，$X^*$ 中任何闭球是弱*紧的.

在本书中，一些较精细的最优化结果依赖于自反空间的特殊性质，为此要用到：

1.1.6 定理 自反空间 X 有以下性质：

(i) 每个 $u \in X^*$ 在 $S(X)$（X 中的单位球面，下文概如此）上

取得最大值$|u|$.

(ii) X 的闭子空间及 X^* 亦为自反空间.

(iii) 任给 $A \subset X$,A 弱紧 $\Leftrightarrow$ A 有界且弱闭,A 有界且闭凸 $\Rightarrow$ A 弱紧.

(iv) X 中任何有界序列含弱收敛子列.

(v) 若 $A \subset X$ 有界,则 $x \in \overline{A}^W \Leftrightarrow$ 存在序列$\{x_n\} \subset A$,使得 $x_n \rightharpoonup x$.

1.1.7 开映射定理 若 $A \in L(X,Y)$,$R(A) = Y$,则 A 是开映射.因此,若 $A: X \to Y$ 是连续线性同构的,则它必为拓扑同构,即 $A^{-1} \in L(Y,X)$.

任给 $D \subset X$,$U \subset X^*$,约定

$$D^\perp = \{u \in X^* : u|D = 0\},$$

$$U^\perp = \{x \in X : u(x) = 0 (\forall u \in U)\}.$$

对于 $A \in L(X,Y)$,以下定理综合了涉及 $N(A)$与 $R(A)$的一系列结论.

1.1.8 定理 设 $A \in L(X,Y)$,则成立:

(i) $N(A) = R(A^*)^\perp$,$N(A^*) = R(A)^\perp$,$N(A^*)^\perp = \overline{R(A)}$,$N(A)^\perp = R(A)^*$的弱*闭包.

(ii)(闭值域定理) $R(A)$为闭集 $\Leftrightarrow$ $R(A^*)$为闭集 $\Leftrightarrow$ $R(A^*) = N(A)^\perp \Leftrightarrow R(A) = N(A^*)^\perp$.

(iii)(满射定理) $R(A) = Y \Leftrightarrow A^*$有有界逆,$R(A^*) = X^*$ $\Leftrightarrow$ A 有有界逆.

本书中有少数结果依赖于空间的一定凸性.若球面 $S(X)$不含多于一点的线段,则说 X 严格凸;若 $\forall x \in S(X)$,$\forall \varepsilon > 0$,$\exists \delta = \delta(x,\varepsilon) > 0$,$\forall y \in S(X)$,使得

$$|x - y| \geqslant \varepsilon \Rightarrow |x + y| \leqslant 2(1 - \delta),$$

则说 X 局部一致凸;若上述的 $\delta(x,\varepsilon)$与 x 无关,则说 X **一致凸**.直接看出,X 是 Hilbert 空间 $\Rightarrow$ X 一致凸 $\Rightarrow$ X 局部一致凸 $\Rightarrow$ X 严格凸.更深入的结果有:

1.1.9 定理 (i) X 严格凸 $\Leftrightarrow \forall u \in X^* \setminus \{0\}$：至多有一点 $x \in S(X)$ 使 $u(x) = \|u\|$；若 X^* 严格凸，$0 \neq x \in X$，则有唯一 $u \in S(X^*)$ 使 $u(x) = \|x\|$.

(ii) 若 X 局部一致凸，在 X 中 $x_n \rightharpoonup x$ 且 $\|x_n\| \to \|x\|$，则 $x_n \to x$.

(iii) 若 X 一致凸，则 X 为自反空间.

(iv) 若 X 自反，则可改赋等价范数，使 X 与 X^* 同时成为局部一致凸空间.

对上述内容的详细讨论可参看文献[236,238].

§1.2 微分理论

以下设 $D \subset X, f: D \to \mathbf{R}, g: D \to Y$. 本书将用到的连续性概念可概述如下.

1.2.1 定义 1° 若 $x_n \to x$ [$x_n \rightharpoonup x$] 恒蕴涵

$$f(x) \leqslant \liminf_n f(x_n),$$

则说 f 下半连续[弱下半连续]，缩写作 lsc [wlsc]. 若 $-f$ 为 lsc [wlsc]，则说 f 上半连续[弱上半连续]，缩写作 usc [wusc].

2° 若 $x_n \rightharpoonup x \Rightarrow g(x_n) \to g(x)$，则说 g 在 x 处强连续. 若 $Y = Z^*$，$\forall a \in X$，当 $t \downarrow 0$ 时 $g(x+ta) \overset{*}{\rightharpoonup} g(x)$，则说 g 在 x 处半连续；若 $x_n \to x \Rightarrow g(x_n) \overset{*}{\rightharpoonup} g(x)$，则说 g 在 x 处次连续.

3° 令 $\mathrm{Lip}g = \sup_{x \neq y} \|g(x) - g(y)\| / \|x - y\|$ (映射 g 的 **Lipschitz 模数**). 若 $\mathrm{Lip}\, g < \infty$，则称 g 为 Lipschitz 连续，记作 $g \in \mathrm{Lip}$. 若 $\forall x \in D$，存在 x 的邻域 U，使 $g|U \in \mathrm{Lip}$，则说 g 是局部 Lipschitz 的，简称为**局部 Lip**，记作 $g \in C^{1-0}$.

本书将用到的导数概念可概述如下.

1.2.2 定义 1° 令 $\varphi(t) = g(x+ta)$. 分别称 $\varphi'_+(0), \varphi'_-(0)$ 与 $\varphi'(0)$(假定它们存在)为 g 在 x 沿 a 的右侧方向导数、左侧方向

导数与方向导数，依次记为 $D_+g(x,a)$，$D_-g(x,a)$ 与 $Dg(x,a)$.

2° 若 $\forall a \in X$，$Dg(x,a)$ 存在，且 $Dg(x,\cdot) \in L(X,Y)$，则说 g 在 x 处 **G-可微**，并称 $Dg(x) \triangleq Dg(x,\cdot)$ 为 g 在 x 的 **G-导数**. 若 $Df(x)$ 存在，则 $Df(x) \in X^*$，通常称 $Df(x)$ 为 f 在 x 的**梯度**，并记作 $\nabla f(x)$.

3° 若 $\varphi(t) = g(x+ta)$，$\varphi^{(n)}(0)$ 存在，则称它为 g 在 x 沿 a 的 n 阶方向导数，记作 $D^n g(x,a)$.

4° 令 $\Delta g(x,z) = g(x+z) - g(x)$. 若存在 $A \in L(X,Y)$，使得 $\Delta g(x,z) = Az + o(z)$ ($\Leftrightarrow |\Delta g(x,z) - Az| = o(|z|)$)，则说 g 在 x 可微，称 A 为 g 在 x 的 **Fréchet 导数**，简称为导数，记作 $g'(x)$. 归纳地可定义 g 的 r 阶导数 $g^{(r)}(x)$. 若 $g^{(r)}(x)$ 在 D 上处处存在且连续，则称 g 为 C^r 映射.

5° 设 $\varphi: \Omega \subset X \times Y \to Z$. 给定 y，若 $\varphi(\cdot,y)$ 可微，则称其导数为 $\varphi(x,y)$ 对 x 的偏导数，记作 $\varphi_x(x,y)$. $\varphi_y(x,y)$ 的意义仿此. 高阶偏导数类似地定义. 若 φ 的所有 r 阶偏导数存在且连续，则称 φ 为 C^r 映射.

1.2.3 中值定理 设 $[x,y] \subset D$ 是一线段.

(i) 若 f 在 $[x,y]$ 上 G-可微，则存在 $z \in [x,y]$，使

$$f(x) - f(y) = \langle \nabla f(z), x - y \rangle.$$

(ii) 若 g 在 $[x,y]$ 上 G-可微，则

$$|g(x) - g(y)| \leqslant \sup_{z \in [x,y]} |Dg(z)|\,|x-y|.$$

1.2.4 Taylor 公式 设 $[x,x+z] \subset D$ 是一线段.

(i) 若 f 在 $[x,x+z]$ 上 $n+1$ 次可微，则存在 $\theta \in (0,1)$，使

$$f(x+z) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x) z^k + \frac{1}{(n+1)!} f^{(n+1)}(x+\theta z) z^{n+1},$$

其中已约定 $f^{(0)}(x)z^0 = f(x)$.

(ii) 若 $g \in C^{n+1}$，则

$$g(x+z) - \sum_{k=0}^{n} \frac{1}{k!} g^{(k)}(x) z^k$$

$$= \frac{1}{n!}\int_0^1 (1-t)^n g^{(n+1)}(x+tz)z^{n+1}\mathrm{d}t = o(|z|^{n+1}).$$

1.2.5 隐函数定理 设 $\varphi: \Omega \subset X\times Y \to Z$ 为 C^r 函数，$r \geqslant 1$，$\varphi(x_0,y_0) = 0$，$\varphi_y(x_0,y_0): Y \to Z$ 是一同构. 则在点 (x_0,y_0) 邻近方程 $\varphi(x,y) = 0$ 有唯一解 $y = g(x)$，使 $y_0 = g(x_0)$，且 $g \in C^r$.

由隐函数定理可推出反函数定理：若 $g \in C^r$，$r \geqslant 1$，$g'(x_0)$：$X \to Y$ 是一同构，则在 $y_0 = g(x_0)$ 邻近 $y = g(x)$ 有反函数 $x = g^{-1}(y)$，且 $g^{-1} \in C^r$.

1.2.6 Lusternik 定理 设 $g \in C^1$，$R(g'(x_0)) = Y$，则当 $z \in X$，$|z|$ 充分小时方程 $g(x_0 + y + z) = g(x_0) + g'(x_0)z$ 有解 $y = y(z) = o(z)(z \to 0)$. 特别，若 $z \in N(g'(x_0))$，则当 $t \to 0$ 时方程 $g(x_0 + y + tz) = g(x_0)$ 有解 $y = y(t) = o(t)$.

设 $A \in L(X,Y)$. 若有拓扑直和分解 $X = N(A) \oplus V$，则说 A 是核裂的. 当 $\dim X < \infty$ 时 A 总是核裂的.

1.2.7 定义 若 $g \in C^r(r \geqslant 1)$，$R(g'(x_0)) = Y$ 且 $g'(x_0)$ 是核裂的，则说 g 在 x_0 处为浸满或 C^r 浸满.

若 $g = (g_i): D \subset \mathbf{R}^n \to \mathbf{R}^m$ 是 C^r 映射，$r \geqslant 1$，$x_0 \in D$，则 g 在 x_0 为浸满 $\Leftrightarrow \{\nabla g_i(x_0)\}$ 线性无关.

1.2.8 定理 设 g 在 x_0 为 C^r 浸满，$A = g'(x_0)$，$M = \{x \in X : g(x) = g(x_0)\}$. 则存在一个从 $N(A)$ 的某 0-邻域到 x_0 在 M 上的某邻域的同胚 φ，$\varphi \in C^r$，且 $\varphi(z) = x_0 + z + o(z)(z \to 0)$.

证 设 $X = X_1 \oplus X_2$，其中 $X_1 = N(A)$. 令

$$\Phi(x_1,x_2) = g(x_0 + x_1 + x_2) - g(x_0) \quad (x_i \in X_i, i = 1,2),$$

则 $\Phi \in C^r$，$\Phi_{x_2}(0,0) = A|X_2 : X_2 \to Y$ 为同构. 于是由 1.2.5 可从方程 $\Phi(x_1,x_2) = 0$ 局部地解出 $x_2 = \psi(x_1)$，使得 $\psi(0) = 0$，$\psi \in C^r$，$\psi'(0) = 0$. 令 $\varphi(z) = x_0 + z + \psi(z)$，则 φ 即合乎定理所求. □

§1.3 多值映射

任给非空集 T,S,称任何映射 $F:T\to 2^S$ 为从 T 到 S 的多值映射或集值映射,称 $\mathrm{Gr}F \overset{\text{def}}{=} \{(x,y):y\in Fx\}$ 为 F 的图形.显然,F 与 $\mathrm{Gr}\,F$ 相互唯一确定,因此常将二者视为等同.任给 $A\subset T$,令 $FA=\bigcup\limits_{x\in A}Fx$.称 $R(F)\overset{\text{def}}{=}FT$ 为 F 的值域,称 $D(F)\overset{\text{def}}{=}\{x\in T:Fx\neq\varnothing\}$ 为 F 的定义域.当写出 $F:D\subset T\to 2^S$ 时,认定 $D=D(F)$ 且 $D\neq\varnothing$. F 的逆 $F^{-1}:S\to 2^T$ 定义为 $F^{-1}y=\{x:y\in Fx\}$.显然 $D(F^{-1})=R(F)$,$R(F^{-1})=D(F)$.所宜注意者,多值映射的逆恒可定义,这是多值映射的主要优点之一.对任给 $A\subset T$,F 在 A 上的限制 $F|A$ 定义为 $(F|A)(x)=Fx\,(x\in A)$.若 $F:T\to 2^S$ 与 $G:S\to 2^W$ 是两个多值映射,则 G 与 F 的复合映射 $G\circ F$ 定义为 $(G\circ F)(x)=G(Fx)\,(x\in T)$.易见 $R(G\circ F)\subset R(G)$,$D(G\circ F)\subset D(F)$,$(G\circ F)^{-1}=F^{-1}\circ G^{-1}$.不致混淆时,可将 $G\circ F$ 写作 GF.

关于多值映射的进一步的概念涉及空间的一定结构.

1.3.1 定义 设 $F:X\to 2^Y$.

1° 若对任给有界集 $A\subset X$,FA 必有界,则说 F 有界.若 $\forall x\in X$,存在 x 的邻域 V,使 FV 有界,则说 F 局部有界.

2° 若 $\mathrm{Gr}\,F$ 是 $X\times Y$ 的闭子集,则说 F 是闭映射.若 $Y=Z^*$,当 X 与 Z^* 中分别用强拓扑与弱*拓扑时 $\mathrm{Gr}\,F$ 为闭集,则说 F 是 sw* 闭映射. sw 闭映射仿此.

3° 若 $\forall\varepsilon>0$,存在 x 的邻域 V,使得 $FV\subset Fx+B_\varepsilon(0)$,则说 F 在 x 处上半连续(usc).若 $Y=Z^*$,任给 Fx 的弱*邻域 V,存在 $\delta>0$,使得 $FB_\delta(x)\subset V$,则说 F 在 x 处为 sw* 上半连续(sw*-usc).若 F 在每点 $x\in D(F)$ 为 usc[sw*-usc],则说 F 为 usc[sw*-usc].

对于单值映射，以上所定义的 usc 与 sw*-usc 分别为连续与次连续.

1.3.2 定义 设 $F: X \to 2^{X^*}$. 若 $\forall x, y \in D(F)$，有 $\langle Fx - Fy, x-y\rangle \geqslant 0$，则说 F 是**单调映射**. 若当 $x, y \in D(F)$，$x \neq y$ 时 $\langle Fx - Fy, x-y\rangle > 0$，则说 F 是**严格单调**的. 若有 $\lambda > 0$，使得 $\langle Fx - Fy, x-y\rangle \geqslant \lambda |x-y|^2 (\forall x, y \in D(F))$，则说 F 是**强单调**的. 若 F 是单调的，且当 $\langle Fx - v, x-y\rangle \geqslant 0\ (\forall x \in D(F))$ 时 $v \in Fy$，则说 F 是**极大单调的**.

1.3.3 例 任给 $x \in X$，令 $Jx = \{u \in X^* : u(x) = |u|^2 = |x|^2\}$，用 1.1.1 易指明 $Jx \neq \varnothing$. 称如此定义的映射 $J: X \to 2^{X^*}$ 为 X 的**对偶映射**. 若 $u \in Jx$，$v \in Jy$，则

$$\begin{aligned}\langle u - v, x - y\rangle &= |x|^2 - u(y) - v(x) + |y|^2 \\ &\geqslant |x|^2 - 2|x||y| + |y|^2 \geqslant 0.\end{aligned}$$

可见 J 是单调的. 其次设 $y \in X$，$v \in X^*$ 满足 $\langle Jx - v, x - y\rangle \geqslant 0 (\forall x \in X)$，则分别取 $x = y + tz$ 与 $x = (1 \pm t)y (t > 0)$ 得

$$v(z) \leqslant \langle J(y - tz), z\rangle,$$

$v(y) = |y|^2$，由此推出 $v \in Jy$. 可见 J 是极大单调的.

1.3.4 命题 设 $F: D \subset X \to 2^{X^*}$ 是单调映射，则 F 在 D 内部局部有界；若 F 是单值的且半连续，则 F 在 D 内部次连续.

1.3.5 命题 (i) 若 $F: D \subset X \to 2^{X^*}$ 是极大单调映射，则 F 是 sw* 闭与 ws 闭的，每个 Fx 为弱* 闭凸集；若 X 自反，则 F 在 D 内部是 sw-usc 的.

(ii) 设 $F: X \to X^*$ 是单调映射. 若 F 半连续，则 F 极大单调；若 X 自反，则 F 极大单调 $\Leftrightarrow$ F 次连续.

(iii) 若 $A: X \to X^*$ 是单调的线性映射，则 A 极大单调且有界(即 $A \in L(X, X^*)$).

1.3.6 定理 设 X 自反，$F: X \to 2^{X^*}$ 极大单调且

$$\lim_{x \in D(F), |x| \to \infty} d(0, Fx) = \infty$$

(当 $D(F)$ 有界时认定此条件自动满足),则 $R(F) = X^*$. 特别,若 $F: X \to X^*$ 极大单调且 $\lim\limits_{|x| \to \infty} |Fx| = \infty$,则 $R(F) = X^*$;若再假定 F 严格单调,则 F 为双射.

1.3.7 推论 设 X 自反,$A: X \to X^*$ 是强单调的线性映射,则 $A^{-1} \in L(X^*, X)$ 且 A^{-1} 亦强单调.

证 由 1.3.5,$A \in L(X, X^*)$ 且极大单调. A 强单调意味着 $\exists \beta > 0, \forall x \in X: \langle Ax, x \rangle \geqslant \beta |x|^2$. 这推出

$$|Ax| \geqslant \beta |x|, \quad x \in X.$$

因此,当 $|x| > 0$ 时 $|Ax| > 0$,可见 A 为单射;当 $|x| \to \infty$ 时 $|Ax| \to \infty$,可见 A 为满射(1.3.6),从而为双射. 由 1.1.7,$A^{-1} \in L(X^*, X)$. □

关于本节的内容可参考文献[93,238].

§1.4 锥与对偶锥

在本书中,锥是一个基本的几何工具,它用于定义序或描述具一定性质的方向. 在最优化理论的研究中,运用锥的技巧是必不可少的.

1.4.1 定义 设 $\varnothing \neq K \subset Y$. 若 $\forall \alpha > 0: \alpha K \subset K$,则称 K 为**锥**;当 $\dim \operatorname{span} K < \infty$ 时,说 K 是有限维锥.

若 $K \subset Y$ 是锥,则 K 是**凸锥** $\Leftrightarrow K + K \subset K$,$K$ 是闭锥 $\Rightarrow 0 \in K$. 若 K 是闭凸锥,定义 $x \leqslant y \Leftrightarrow y - x \in K \Leftrightarrow y \geqslant x (\forall x, y \in Y)$,则 $\leqslant$ 是 Y 中的**拟序**. 本书中,总以 Y_+ 记 Y 中某个给定的闭凸锥,且当未作说明时在 Y 中使用由 Y_+ 定义的拟序 $\leqslant$,并约定 $x < y \Leftrightarrow y - x \in Y_+ \setminus \{0\} \Leftrightarrow y > x$,$x \ll y \Leftrightarrow y - x \in Y_+^\circ \Leftrightarrow y \gg x (x, y \in Y)$,$Y_- = -Y_+$. 当 $Y_+ \cap Y_- = \{0\}$ 时称 Y_+ 为**序锥**. Y_+ 是序锥 $\Leftrightarrow$ 它定义的拟序 $\leqslant$ 为半序. 在 $\mathbf{R}^n$ 中总使用由锥 $\mathbf{R}^n_+$ 定义的**自然向量序** $\leqslant$,即 $x \leqslant y \Leftrightarrow y - x \in \mathbf{R}^n_+ \Leftrightarrow x_i \leqslant y_i$ (1

$\leqslant i \leqslant n, x, y \in \mathbf{R}^n$).

任给 $A \subset X, A \neq \varnothing$,称 cone $A \stackrel{\text{def}}{=} \bigcup_{\alpha > 0} \alpha A$ 为 A 生成的锥,记 $\overline{\text{cone}}\ A = \overline{\text{cone}\ A}$;称 $A^* \stackrel{\text{def}}{=} \{u \in X^* : u(A) \geqslant 0\}$为 A **的对偶锥**,约定$\varnothing^* = X^*$. 容易验证以下简单结论:A^* 总是 X^* 中的弱*闭凸锥;$A \subset B \Rightarrow A^* \supset B^*$;$A \subset A^{**}$;$A^* = (\overline{\text{co}}\ \overline{\text{cone}} A)^*$;$A$ 是子空间 $\Rightarrow A^* = A^{\perp}$;$A$ 凸且 $A^{\circ} \neq \varnothing \Rightarrow A^* = (A^{\circ})^*$. 注意 X 本身的对偶锥是$\{0\}$(不宜将它写作 X^*,否则与对偶空间混淆!).

1.4.2 命题 设 $K \subset Y$ 是锥.

(i) 若 $\lambda \in K^*$,则 $\inf \lambda(K) = 0$;若 $\lambda \in Y^* \backslash K^*$,则 $\inf \lambda(K) = -\infty$. 因此,若 $\lambda(K) \geqslant r > -\infty$, 则 $\lambda \in K^*$ 且 $r \leqslant 0$; 但若 $\lambda(K) \leqslant r < \infty$,则 $\lambda \in -K^*$ 且 $r \geqslant 0$.

(ii) 若 $y \in K^{\circ}$,$0 \neq \lambda \in K^*$,则 $\lambda(y) > 0$;当 $\lambda \in K^*$, $\|\lambda\| \to \infty$时 $\lambda(y) \to \infty$.

(iii) 若 K 闭凸且 $K^{\circ} \neq \varnothing, y \in \partial K$,则有 $0 \neq \lambda \in K^*$: $\lambda(y) = 0$.

1.4.3 定理 $Y_+ = Y \cap Y_+^{**}$;当 Y 自反时,$Y_+ = Y_+^{**}$.

以上结论今后将反复利用,它表明 Y_+ 与 Y_+^{**} 的关系有如 Y 与 Y^{**} 的关系. 特别,利用 1.4.3 并结合 1.4.2,得出以下推论.

1.4.4 推论 任给 $y \in Y$,有 $y \geqslant 0 \Leftrightarrow \langle Y_+^*, y \rangle \geqslant 0$;若 $Y_+^{\circ} \neq \varnothing$,则 $y \gg 0 \Leftrightarrow \langle Y_+^* \backslash \{0\}, y \rangle > 0$.

1.4.4 是判定“向量不等式”的基本依据:任给 $y, z \in Y$,有 $y \geqslant z \Leftrightarrow \lambda(y) \geqslant \lambda(z) (\forall \lambda \in Y_+^*)$;当 $Y_+^{\circ} \neq \varnothing$时,$y \gg z \Leftrightarrow \lambda(y) > \lambda(z) (\forall \lambda \in Y_+^* \backslash \{0\})$.

1.4.5 Krein 定理 设 $A \subset Y$ 为子空间, $K \subset Y$ 为凸锥, 且 $A \cap K^{\circ} \neq \varnothing$,$\lambda$是 A 上的连续线性泛函,$\lambda(A \cap K) \geqslant 0$. 则存在 $v \in K^*$,使 $v|A = \lambda$.

在锥概念对最优化理论的应用中,最基本的操作是求对偶锥.

为此,需用到关于$\left(\sum K_i\right)^*$,$(\bigcap K_i)^*$的计算公式. 首先,有以下简单结果.

1.4.6 命题 设$K_i \subset Y(1 \leqslant i \leqslant n)$非空且是锥,至多一个例外,则$\left(\sum K_i\right)^* = \bigcap K_i^*$.

形式上类似但更有意义的问题是:若$K_i \subset Y(1 \leqslant i \leqslant n)$是锥,是否成立

$$(\bigcap K_i)^* = \sum K_i^* ? \tag{1}$$

包含$(\bigcap K_i)^* \supset \sum K_i^*$是平凡的,但相反的包含依赖于一定附加条件,且其证明是非平凡的.

1.4.7 定理 设$K_i \subset Y(1 \leqslant i \leqslant n)$是锥,则以下每个条件推出等式(1):

(i) K_i是闭凸锥且$\sum K_i^*$弱*闭.

(ii) K_i是开凸锥且$\bigcap K_i \neq \varnothing$.

(iii) $n = 2$,K_1是子空间,K_2是凸锥且$K_1 \cap K_2^\circ \neq \varnothing$.

证 只要对任给$\lambda \in (\bigcap K_i)^*$,证$\lambda \in \sum K_i^*$. 若条件(i)满足而$\lambda \overline{\in} \sum K_i^*$,则由1.1.3有$x \in X$:$\lambda(x) < \langle \sum K_i^*, x\rangle$,这推出$\lambda(x) < 0$,$\langle K_i^*, x\rangle \geqslant 0$,从而$x \in \bigcap K_i$(1.4.4),这矛盾于$\lambda \in (\bigcap K_i)^*$. 若条件(ii)满足,令$M = \{(y, y, \cdots, y) : y \in Y\}$,$K = \prod K_i$,$f(y, y, \cdots, y) = \lambda(y)$,在$Y^n$中对子空间$M$与开凸锥$K$应用1.4.5,得$\varphi = (\lambda_i) \in K^* = \prod K_i^*$:$\varphi|M = f$. 这推出$\lambda = \sum \lambda_i \in \sum K_i^*$. 若条件(iii)满足,则对$\lambda|K_1$应用1.4.5得$v \in K_2^*$:$v|K_1 = \lambda|K_1$. 于是

$$\lambda = (\lambda - v) + v \in K_1^\perp + K_2^* = K_1^* + K_2^*. \qquad \square$$

关于1.4.2~1.4.7可参考文献[93,238];进一步的结果见§3.1.

经常需要判定X^*中一凸集为弱*闭(如见1.4.7(i)),当X自反时,这归于判定强闭(1.1.3),为此可利用以下结果.

1.4.8 定理　设 $A, B \subset X$ 闭凸，A 局部紧且 $C_A \cap C_B = \{0\}$，其中 $C_A = \{x : x + A \subset A\}$，则 $B - A$ 为闭集. 由此推出：若 A 是 X 的闭子空间，$B \subset X$ 是闭凸锥，$A \cap B = \{0\}$，A 或 B 是有限维的，则 $A + B$ 是闭的.

注意，即使在有限维空间中，两个闭凸锥之和也未必是闭的. 例如，在 $\mathbf{R}^3$ 中 $A = 0 \times \mathbf{R} \times 0$ 与 $B = \{(x, y, z) : x^2 + z^2 \leqslant 2yz, y \geqslant 0\}$ 分别为子空间与闭凸锥，但 $A + B = A \cup (\mathbf{R}^2 \times (0, \infty))$ 是非闭的.

参考文献：[23, 93, 116, 177, 238].

§1.5 凸　函　数

涉及凸函数的最优化问题往往有较好的结论. 因此，各种意义下的凸函数在本书中有基本的重要性. 以下设 $D \subset X, f : D \to \mathbf{R}, g : D \to Y, G : D \to 2^Y$；$Y$ 中已由闭凸锥 Y_+ 导入序 $\leqslant$；$J = [0, 1]$.

1.5.1 定义　设 D 是凸集. 若 $\forall x, y \in D, t \in J$，有

$$tGx + t'Gy \subset G(tx + t'y) + Y_+$$

($t' = 1 - t$，今后概如此)，则称 G 为 Y_+-凸映射；若 $\forall x, y \in D$，$x \neq y, t \in (0, 1)$，有

$$tGx + t'Gy \subset G(tx + t'y) + (Y_+ \backslash 0),$$

则说 G 是严格 Y_+-凸的. f 是凸[**严格凸**]的意指它是 $\mathbf{R}_+$-凸[严格 $\mathbf{R}_+$-凸]的.

关于**凸映射**的一些等价刻画如下.

1° 令 $\operatorname{epi} G = \{(x, y) : y \in Gx + Y_+\}$（称之为 G 的**上图**）. 则 G 是 Y_+-凸映射 $\Leftrightarrow$ $\operatorname{epi} G$ 是凸集. 注意

$$\operatorname{epi} g = \{(x, y) : y \geqslant g(x)\}.$$

2° G 是 Y_+-凸映射 $\Leftrightarrow$ 对任何凸组合 $x = \sum t_i x_i \in D$ ($t_i \geqslant 0, \sum t_i = 1, x_i \in D, 1 \leqslant i \leqslant n$)，有 $\sum t_i Gx_i \subset G(x) + Y_+$. 特

别，g 是 Y_+-凸映射 $\Leftrightarrow$ 对上述的 $x=\sum t_i x_i$ 有 $g(x)\leqslant\sum t_i g(x_i)$；$g$ 是严格 Y_+-凸的 $\Leftrightarrow$ 当 $x,y\in D,x\neq y,t\in(0,1)$ 时 $g(tx+t'y)<tg(x)+t'g(y)$.

3° g 是 Y_+-凸映射 $\Leftrightarrow \forall\lambda\in Y_+^*:\lambda\circ g$ 是凸函数. 若 $g=(g_i):D\to\mathbf{R}^n$，则 g 是凸的 $\Leftrightarrow$ 每个 g_i 是凸的.

关于凸函数的连续性有一些很强的结论.

1.5.2 定理 设 f 是凸开集 D 内的凸函数. 则以下每个条件推出 $f\in C^{1-0}$：(i) f 在某个球 $B(\subset D)$ 内上有界；(ii) f 在某个球内为 lsc；(iii) f 在 D 内某点连续；(iv) $\dim X<\infty$. 若 f 在 D 内为 lsc 或 G-可微，则 f 必为 wlsc.

凸函数的微分性质可综合于下.

1.5.3 定理 设 D 是凸开集.

(i) 若 f 是凸函数，则 $D_+f(x,z)$ 恒存在，且

$D_-f(x,z)\leqslant D_+f(x,z)\leqslant\Delta f(x,z)(x,x+z\in D)$；

若 $\dim X<\infty$，$\nabla f(x)$ 存在，则 $f'(x)=\nabla f(x)$.

(ii) 设 f 在 D 上是 G-可微的，则 f 凸 $\Leftrightarrow$ 当 $x,x+z\in D$ 时 $\langle\nabla f(x),z\rangle\leqslant\Delta f(x,z)\Leftrightarrow\nabla f$ 是单调映射；f 严格凸 $\Leftrightarrow$ 当 $x,x+z\in D,z\neq0$ 时 $\langle\nabla f(x),z\rangle<\Delta f(x,z)\Leftrightarrow\nabla f$ 严格单调；若 f 凸，则 ∇f 半连续，当 X 自反或可分时 ∇f 次连续，当 $\dim X<\infty$ 时 $f\in C^1$.

(iii) 若 $D^2f(x,z)$ 存在，则 f 凸 $\Leftrightarrow\forall x\in D,z\in X:D^2f(x,z)\geqslant0$；$D^2f(x,z)>0(\forall x\in D,0\neq z\in X)\Rightarrow f$ 严格凸.

(iv) 若 $\exists\lambda>0,\forall x\in D,\forall z\in X:D^2f(x,z)\geqslant\lambda|z|^2$，则 ∇f 强单调. 反之，若 ∇f 强单调，则有 $\lambda>0$，当 $x,x+z\in D$ 时

$$\langle\nabla f(x),z\rangle\leqslant\Delta f(x,z)-(\lambda/2)|z|^2.$$

1.5.3 中的某些结论亦适用于向量函数. 例如，若 g 是 G 可微的，则 g 凸 $\Rightarrow$ 当 $x,x+z\in D$ 时 $Dg(x)z\leqslant\Delta g(x,z)$；其次若 $D^2g(x,z)$ 恒存在，则 g 凸 $\Leftrightarrow\forall x\in D,\forall z\in X:D^2g(x,z)\geqslant0$.

1.5.4 定理 若 f 凸且为 lsc,则存在 $\alpha,\beta \geqslant 0$,使

$$f(x) \geqslant -\alpha|x| - \beta \ (\forall x \in D).$$

关于 1.5.2～1.5.4 的证明可参考[93].

在凸函数对最优化理论的应用中,往往只用到凸函数的多个性质中的某一特定性质;倘用该性质界定一类函数,则这类函数较凸函数为广(因而称为广义凸函数),而在最优化理论中可起到类似于凸函数的作用.因此,在最优化理论中,广义凸函数的使用极为广泛,广义凸函数的种类也极为多样.下面仅对某些常用的广义凸函数作一综合.首先考虑可微函数,以下的广义凸性定义是从 1.5.3(ii)引伸出来的.

1.5.5 定义 设 $F: D\times X^* \to \mathbf{R}$,$F(x,\cdot)$是次线性的;$r \in \mathbf{R}$;$\theta: D\times D \to \mathbf{R}_+$,$\theta(x,y) = 0 \Leftrightarrow x = y$;$x \in D$,$\nabla f(x)$存在.

(i) 若 $\forall y \in D\backslash\{x\}$:$f(y)-f(x) \geqslant F(y,\nabla f(x))+r\theta(x,y)$[$f(y)-f(x) > F(y,\nabla f(x))+r\theta(x,y)$],则说 f 在 x 关于 D 为(F,r,θ)-凸[(F,r,θ)-严格凸].

(ii) 若 $\forall y \in D$: $F(y,\nabla f(x)) + r\theta(x,y) > 0 \Rightarrow f(y) > f(x)$,则说 f 在 x 关于 D 为(F,r,θ)-拟凸.

(iii) 若 $\forall y \in D\backslash\{\bar{x}\}$:$f(y,\nabla f(x))+r\theta(x,y) \geqslant 0 \Rightarrow f(y) \geqslant f(x)$[$f(y) > f(x)$],则说 f 在 x 关于 D 为(F,r,θ)-伪凸[(F,r,θ)-严格伪凸](参考[171]).

以上定义可扩张到整个 D 上:设 $F: D\times D\times X^* \to \mathbf{R}$,$F(x,y,\cdot)$是次线性的;$f$ 在 D 上 G-可微.若 $\forall x,y \in D$:$f(y)-f(x) \geqslant F(y,x,\nabla f(x))+r\theta(x,y)$,则说 f 在 D 上为(F,r,θ)-凸.余仿此.若 $r > 0$[$r < 0$],则称(F,r,θ)-凸为强(F,θ)-凸[弱(F,θ)-凸];若 $r = 0$,则(F,r,θ)-凸就称为(F,θ)-凸.其余类推.

定义 1.5.5 概括了多种特殊情况.

1° 若 $\theta(x,y) = |x-y|$,则称(F,r,θ)-凸为(F,r)-凸;余类推.

2° 若 $F(x,\xi) = \langle\xi,\eta(x)\rangle$,$\eta: D \to X$ 是给定的,则称(F,r)-

凸为(η,r)-凸. f 在 D 上(η,r)-凸意味着存在 $\eta: D\times D\to X$, $r\in\mathbf{R}$, $\forall x,y\in D$: $f(y)-f(x)\geqslant\langle\nabla f(x),\eta(x,y)\rangle+r|x-y|$.

3° 若 $r=0$,则称(η,r)-凸为 η-凸,余类推. [83]最早考虑"η-凸性",当时称为不变凸(invex).

4° 若 $\eta(x,y)=y-x$,则 η-凸与 η-严格凸分别为通常的凸与严格凸,而"η-拟凸"就称为拟凸;"η 伪凸"与"η 严格伪凸"仿此. 为明确起见,分别给出 f 在 x 关于 D 凸、严格凸、拟凸、伪凸、严格伪凸的直接刻画如下:

$$\forall y\in D: f(y)-f(x)\geqslant\langle\nabla f(x),y-x\rangle; \tag{1}$$

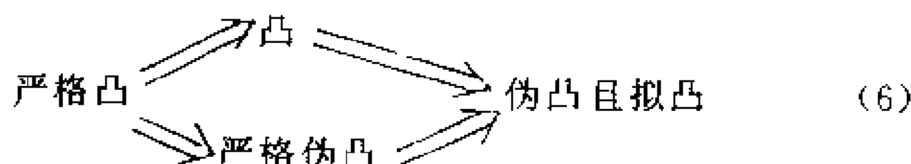
$$\forall y\in D\backslash\{x\}: f(y)-f(x)>\langle\nabla f(x),y-x\rangle; \tag{2}$$

$$\forall y\in D: \langle\nabla f(x),y-x\rangle>0\Rightarrow f(y)>f(x); \tag{3}$$

$$\forall y\in D: \langle\nabla f(x),y-x\rangle\geqslant 0\Rightarrow f(y)\geqslant f(x); \tag{4}$$

$$\forall y\in D\backslash\{x\}: \langle\nabla f(x),y-x\rangle\geqslant 0\Rightarrow f(y)>f(x). \tag{5}$$

直接看出,以上诸凸性有如下关系:



若在每种凸性之前同时加上"η-",或"(η,r)-",或"(F,r)-",则逻辑关系(6)仍然保持.

下面用若干初等例子来说明以上诸概念的一般性及其相互关系.

1.5.6 例 1° 设 $f(x)$在 $\mathbf{R}$ 上可微且 $f'(x)>0$,令 $\eta_\epsilon(x,y)=[f(y)-f(x)+\epsilon]/f'(x)$,则

$$f'(x)\eta_\epsilon(x,y)=f(y)-f(x)+\epsilon.$$

由此可见,当 $\epsilon>0$ 时 f 为 η_ϵ-严格凸;当 $\epsilon=0$ 时,f 为 η_ϵ-凸而非 η_ϵ 严格凸;当 $\epsilon<0$ 时,f 为 η_ϵ-严格伪凸而非 η_ϵ-凸. 因 f 严格增,故当 $y\neq x$ 时,

$$f'(x)(y-x)\geqslant 0\Rightarrow y>x\Rightarrow f(y)>f(x),$$

可见 f 严格伪凸. 取 $f(x) = 2x+\sin x$ 看出 f 不必凸.

2° 设 $f(x) = \sin^3 x(0 \leqslant x \leqslant \pi)$,

$$\eta(x,y) = \cos x(\sin y - \sin x),$$

则

$$f'(x)\eta(x,y) = 3\sin^2 x\cos^2 x(\sin y - \sin x);$$

$$f'(x)\eta(x,y) > 0 \Rightarrow \sin y > \sin x \Rightarrow f(y) > f(x),$$

可见 f 为 η-拟凸. 另一方面, 由 $f(\pi/4) = f(3\pi/4)$ 与 $f'(\pi/4) > 0$ 推出 f 并非拟凸; 由 $f'(\pi/2)\eta(\pi/2, \pi/4) = 0$ 以及 $f(\pi/4) < f(\pi/2)$ 知 f 并非 η-伪凸.

3° 设 $f(x) = x_1 + \sin x_2, x \in D$, 此处

$$D = \{x \in \mathbf{R}_+^2 : x_1^2 + x_2^2 \leqslant 9/4\},$$

$$\eta(x,y) = \left(\frac{\sin y_1 - \sin x_1}{\cos x_1}, \frac{\sin y_2 - \sin x_2}{\cos x_2}\right),$$

则

$$\langle \nabla f(x), \eta(x,y) \rangle = \frac{\sin y_1 - \sin x_1}{\cos x_1} + \sin y_2 - \sin x_2$$

$$= \frac{\cos \tau}{\cos x_1}(y_1 - x_1) + \sin y_2 - \sin x_2 \quad (\tau \in [x_1, y_1])$$

$$\leqslant y_1 - x_1 + \sin y_2 - \sin x_2 = f(y) - f(x),$$

可见 f 为 η-凸. 另一方面, 对 $x = (\pi/6, \pi/3)$ 与 $y = (\pi/3, 0)$ 有

$$\langle \nabla f(x), y - x \rangle = 0, f(y) = \pi/3 < (\pi + 3\sqrt{3})/6 = f(x),$$

可见 f 非伪凸.

4° 设 $f(x) = \cos^2 x(0 < x < \pi/2)$,

$$\eta(x,y) = \sin x(\cos x - \cos y).$$

则

$$f'(x)\eta(x,y) = 2\sin^2 x\cos x(\cos y - \cos x) \geqslant 0;$$

$$f'(x)\eta(x,y) \geqslant 0 \Rightarrow \cos y \geqslant \cos x \Rightarrow f(y) \geqslant f(x),$$

可见 f 为 η-伪凸. 另一方面, 由

$$f'\left(\frac{\pi}{4}\right)\eta\left(\frac{\pi}{4},\frac{\pi}{3}\right)=\frac{\sqrt{2}}{4}-\frac{1}{2}>-\frac{1}{2}$$
$$=f\left(\frac{\pi}{3}\right)-f\left(\frac{\pi}{4}\right)$$

知 f 并非 η-凸.

对于非可微函数,可定义拟凸性如下:

1.5.7 定义 设 D 为凸集. 若 $\forall x,y\in D$、$\forall t\in J$,有

$$f(tx+t'y)\leqslant f(x)\vee f(y)(\overset{\text{def}}{=}\max\{f(x),f(y)\}),$$

则说 f **拟凸**;若 $\forall x,y\in D,t\in(0,1)$,

$$x\neq y\Rightarrow f(tx+t'y)<f(x)\vee f(y),$$

则说 f **强拟凸**;若 $\forall x,y\in D,t\in(0,1)$:

$$f(x)\neq f(y)\Rightarrow f(tx+t'y)<f(x)\vee f(y),$$

则说 f **严格拟凸**.

容易证明:f 拟凸 $\Leftrightarrow\forall x_i\in D,t_i\in\mathbf{R}_+(1\leqslant i\leqslant n)$,当 $\sum t_i=1$ 时 $f(\sum t_ix_i)\leqslant\max f(x_i)\Leftrightarrow\forall\alpha\in\mathbf{R}:X(f\leqslant\alpha)$ 为凸集. 若 f 是 G- 可微的,则 1.5.7 所定义的拟凸与式(3) 所刻画的拟凸一致. 一般地,凸或伪凸 $\Rightarrow$ 拟凸且严格拟凸;严格凸或严格伪凸 $\Rightarrow$ 强拟凸 $\Rightarrow$ 拟凸且严格拟凸.

1.5.8 例 1° 设 f 是实区间上的实函数,则 f 单调 $\Rightarrow f$ 拟凸;f 严格单调 $\Rightarrow f$ 强拟凸. $f(x)=x^3$ 强拟凸,但它在 $x=0$ 处不是伪凸的.

2° 设 $f(x)=0\ (\forall x<0),f(y)=1\ (\forall y\geqslant 0)$,则 f 拟凸而非严格拟凸,若 $f(x)=0\ (\forall x\neq 0),f(0)=1$,则 f 严格拟凸而非拟凸.

1.5.5 与 1.5.7 皆可推广于向量值函数,例如 $g:D\to Y$(关于 Y_+)拟凸意味着:$\forall x_i\in D,\forall y\in Y$,有

$$g(x_i)\leqslant y(i=1,2)\Rightarrow g([x_1,x_2])\leqslant y.$$

但由于向量序与实数序之间有重大差别,形式地定义的广义凸向量值函数可能与相应的广义凸实函数有很不同的性质. 例如,对拟

凸实函数的等价刻画，一般不再适用于拟凸向量值函数. 因此，对于向量值函数通常更适于使用所谓“*”概念：g 是“*拟凸”的 $\Leftrightarrow$ $\forall\lambda \in Y_+^*$：λg 是拟凸的. 其他“*广义凸性”仿此. 可直接看出 * 拟凸 $\Rightarrow$ 拟凸，但其逆不真.

1.5.9 例 设 $g: \mathbf{R} \to \mathbf{R}^2, x \to (\min\{x,1\}, -\min\{x,2\})$，则不难直接验证 g 是拟凸的. 但 g 不是 * 拟凸的：

$$\langle (2,1), g(x) \rangle = 2\min\{x,1\} - \min\{x,2\}$$

$$= \begin{cases} x, & x < 1, \\ 2 - x, & 1 \leqslant x \leqslant 2, \\ 0, & x > 2 \end{cases}$$

不是拟凸的.

参考文献：[93,120,170,171,177,222,238].

§1.6 极　　值

无疑，极值正是本书要系统考察的主题，此处仅给出最基本的定义与结果，以作为更深入讨论的准备. 以下设 $D \subset X, f: D \to \mathbf{R}, \bar{x} \in D$.

1.6.1 定义 若 $f(\bar{x}) = \min F(D)$，则称 $\bar{x}$ 为 f 的极小点（或全局极小点）. 若 $\forall x \in D \setminus \{\bar{x}\}: f(x) > f(\bar{x})$，则称 $\bar{x}$ 为 f 的严格极小点，若存在 $\bar{x}$ 的邻域 N，使 $\bar{x}$ 是 $f|(D \cap N)$ 的极小点[严格极小点]，则称 $\bar{x}$ 为 f 的**局部极小点**[局部严格极小点]. 关于“极大点”的相应概念可类似定义.

关于极值的以下基本结果可看作经典 Weierstrass 定理的推广.

1.6.2 定理 设 X 自反，D 弱闭，f 为 wlsc，$\lim\limits_{x \in D, \|x\| \to \infty} f(x) = \infty$，则 f 有极小点.

证 令 $\alpha = \inf f(D)$. 取 $\{x_n\} \subset D$，使 $f(x_n) \to \alpha$. $\{x_n\}$ 必定有界. 由 X 自反，可设 $x_n \rightharpoonup x_0$. 因 D 弱闭，故 $x_0 \in D$. 于是 $f(x_0)$

$\leqslant \lim\limits_{n} f(x_n) = \alpha$，这推出 $f(x_0) = \alpha$，x_0 即为 f 的极小点. □

注 若 D 有界，则认定条件 $\lim\limits_{x \in D, |x| \to \infty} f(x) = \infty$ 自动满足. 今后在类似情况下都作此理解.

若附加一定凸性条件，则可得较强的结果.

1.6.3 定理 设 D 为凸集.

(i) 若 X 自反，D 为闭集，f 拟凸且 lsc，$\lim\limits_{x \in D, |x| \to \infty} f(x) = \infty$，则 f 有极小点.

(ii) 若 f 强拟凸，则 f 至多有一个极小点.

(iii) 若 f 拟凸，则 f 的极小点之全体是一凸集.

(iv) 若 f 严格拟凸，则 f 的局部极小点是全局极小点.

(v) 设 $D_+ f(\bar{x}, \cdot)$ 存在. 若 $\bar{x}$ 是 f 的极小点，则 $D_+ f(\bar{x}, D - \bar{x}) \geqslant 0$；当 f 凸时其逆亦真.

(vi) 设 $\nabla f(\bar{x})$ 存在. 若 $\bar{x}$ 是 f 的极小点，则 $\langle \nabla f(\bar{x}), D - \bar{x} \rangle \geqslant 0$（当 $\bar{x} \in D^\circ$ 时 $\nabla f(\bar{x}) = 0$）；当 f 在 $\bar{x}$ 关于 D 伪凸时其逆亦真.

证 (i) 由 1.6.2 与 1.1.4 得出. (ii) ~ (iv) 是定义 1.5.7 与 1.6.1 的直接推论. (v)的前一半是明显的. 若 f 凸，则由 1.5.3(i) 有 $f(x) - f(\bar{x}) \geqslant D_+ f(\bar{x}, x - \bar{x})$ ($\forall x \in D$)，由此得出(v)的后一半. (vi)的证明是类似的. □

1.6.4 Ekeland 定理 设 f 在闭集 D 上为 lsc，$\alpha \stackrel{\text{def}}{=} \inf f(D) > -\infty$，$\varepsilon, \delta > 0$，$f(x_0) \leqslant \alpha + \varepsilon$. 则存在 $x \in D \cap \overline{B}_\delta(x_0)$，使得 $f(x) \leqslant f(x_0)$，且 $\forall y \in D \backslash \{x\}$：

$$f(x) < f(y) + (\varepsilon/\delta)|x - y|.$$

特别，有 $x_\varepsilon \in D$，使得 $f(x_\varepsilon) = \min\limits_{y \in D}[f(y) + \varepsilon|x_\varepsilon - y|] \leqslant \alpha + \varepsilon$（称此 x_ε 为 f 的"ε-近似极小点"）.

证 设定理不真. 归纳地取定 $x_n (n \geqslant 0)$ 如下：设已取点 $x_0, x_1, \cdots, x_n$，它们满足

$$|x_i - x_{i-1}| \leqslant \delta\varepsilon^{-1}[f(x_{i-1}) - f(x_i)], \quad i = 1, 2, \cdots, n. \tag{1}$$

令 $S_n=\{y\in D\backslash\{x_n\}: f(x_n)\geqslant f(y)+\delta^{-1}\varepsilon|y-x_n|\}$，则 $S_n\neq\varnothing$（否则 $x=x_n$ 已满足定理要求）．令 $\beta_n=\inf f(S_n)$，则 $f(x_n)>\beta_n\geqslant\alpha$．取 $x_{n+1}\in S_n$，使 $f(x_{n+1})<[\beta_n+f(x_n)]/2$，则(1)亦对 $i=n+1$ 成立．这就得序列 $\{x_n\}\subset D$，它满足

$$|x_{n+1}-x_n|\leqslant\delta\varepsilon^{-1}[f(x_n)-f(x_{n+1})], n=0,1\cdots. \quad (2)$$

式(2)表明 $f(x_n)$ 单调减，于是 $f(x_n)\to\beta\geqslant\alpha$．若 $n>m$，则由式(2)有

$$|x_n-x_m|\leqslant\delta\varepsilon^{-1}[f(x_m)-f(x_n)], \quad (3)$$

可见 $\{x_n\}$ 是 Cauchy 列．设 $x_n\to x\in D$．由式(3)推出

$$|x_n-x_0|\leqslant\delta\varepsilon^{-1}[f(x_0)-f(x_n)]\leqslant\delta\varepsilon^{-1}(\alpha+\varepsilon-\alpha)=\delta,$$

故 $x\in D\cap B_\delta(x_0)$．因 f 为 lsc，故应用式(3)得

$$f(x)\leqslant\varliminf_n f(x_n)\leqslant f(x_m)-\varepsilon\delta^{-1}|x-x_m|(m\geqslant 0). \quad (4)$$

由(4)得 $f(x)\leqslant f(x_0)$．因已设定理不真，故有 $y\in D\backslash\{x\}$，使得 $f(x)\geqslant f(y)+\varepsilon\delta^{-1}|x-y|$．$\forall n\geqslant 1$，由(4)有

$$\begin{aligned}f(x_n)&\geqslant f(x)+\varepsilon\delta^{-1}|x-x_n|\\&\geqslant f(y)+\varepsilon\delta^{-1}|y-x_n|,\end{aligned}$$

可见 $y\in S_n$．因此 $f(y)\geqslant\beta_n>2f(x_{n+1})-f(x_n)$．令 $n\to\infty$ 得 $f(y)\geqslant\beta$．但 $\beta\geqslant f(x)>f(y)$，得出矛盾． □

Ekeland 定理是现代极值理论中最有意义的成果之一，有很多重要应用（参考[47,59]）．

第二章　非光滑分析

在最优化问题中，无论目标函数与约束函数，都可能越出通常可微函数的范围，因而使传统的微分学无以为用. 在这种情况下，本章所讨论的非光滑分析是一个有效的替代工具. 非光滑分析的基本问题是，将传统的微分学方法加以推广，使之能在多少类似的形式下用于一定的不可微函数. 这一设想已得到辉煌的实现. 非光滑分析的结果不仅成功地用于处理非光滑问题，尤其是非光滑最优化问题，而且也用于一些初看起来没有任何非光滑性的问题.

本章限于考虑凸函数与 Lipschitz 函数的"次微分学"，包括次微分的定义、性质、计算方法及某些相关问题.

§2.1　次　微　分

设 $f: X \to \overline{\mathbf{R}} = \mathbf{R} \cup \{\infty\}$，$D = D_f \overset{\text{def}}{=} \{x : f(x) < \infty\} \neq \varnothing$.

2.1.1 定义　设 $x \in D$. 称

$$\partial f(x) \overset{\text{def}}{=} \{u \in X^* : u \leqslant \Delta f(x, \cdot)\}$$

为 f 在 x 处的**次微分**. 若 $\partial f(x) \neq \varnothing$，则说 f 在 x 处**次可微**，且称每个 $u \in \partial f(x)$ 为 f 在 x 处的**次梯度**.

注　**1°** 易验知 $u \in \partial f(x) \Leftrightarrow \{(y,t) : t - f(x) = u(y-x)\}$ 是 epi f 在点 $(x, f(x))$ 处的支承超平面. 这一事实可与导数的切线解释相对照.

2° 设 f 在凸集 D 上次可微. 若 $z = tx + t'y \in [x,y] \subset D$，取 $u \in \partial f(z)$，则

$$\begin{aligned} f(z) &= tf(z) + t'f(z) \\ &\leqslant t[f(x) + u(z-x)] + t'[f(y) + u(z-y)] \\ &= tf(x) + t'f(y), \end{aligned}$$

可见 f 凸. 这一事实表明，次微分概念本质上仅适用于凸函数，尽管定义 2.1.1 本身并不涉及函数的凸性.

2.1.2 例 **1°** 对一凸函数 $\varphi: \mathbf{R} \to \mathbf{R}$，容易验证

$$\partial\varphi(x) = [\varphi'_-(x), \varphi'_+(x)].$$

2° 设 $f(x) = |x|^2/2$ $(x \in X)$，今指明 $\partial f = J$，J 依 1.3.3. 显然 $\partial f(0) = J(0) = \{0\}$. 其次设 $x \neq 0$. 若 $u \in \partial f(x)$，则

$$\begin{aligned} |u||x| &= \sup_{|y| = |x|} u(y) = u(x) \leqslant \mathrm{D}_+ f(x,x) \\ &= |x|^2 = \mathrm{D}_- f(x,x) \leqslant u(x), \end{aligned}$$

这得出 $u \in Jx$. 反之，若 $u \in Jx$，则 $\forall y \in X$：

$$u(y-x) \leqslant |u||y| - |x|^2 \leqslant f(y) - f(x),$$

可见 $u \in \partial f(x)$. 故得 $\partial f(x) = Jx$.

3° 设 $f(x) = |x|$ $(x \in X)$. 若 $x \neq 0$，则

$$\begin{aligned} \partial f(x) &= |x|^{-1}\partial(|x|^2/2) = |x|^{-1}Jx \\ &= \{u \in X^* : |u| = 1, u(x) = |x|\}. \end{aligned}$$

$u \in \partial f(0) \Leftrightarrow \forall x \in X: u(x) \leqslant |x| \Leftrightarrow |u| \leqslant 1$，因此

$$\partial f(0) = \overline{B}_1(0) \subset X^*.$$

4° 任给 $A \subset X$，A 的**指示函数** δ_A 定义为

$$\delta_A(x) = \begin{cases} 0, & x \in A, \\ \infty, & x \in A^c. \end{cases}$$

任给 $x \in A$，有 $\partial\delta_A(x) = (x-A)^*$. 若 A 是锥，则

$$\partial\delta_A(x) = (-A)^* \cap \{x\}^\perp;$$

若 A 是子空间，则 $\partial\delta_A(x) = A^\perp$.

下面给出 f 在一点次可微的必要与充分条件.

2.1.3 命题 设 f 凸，$x \in D$.

(i) 若 f 在 x 处次可微，则 f 在 x 为 lsc.

(ii) $u \in \partial f(x) \Leftrightarrow u \leqslant D_+ f(x, \cdot) \Leftrightarrow D_- f(x, \cdot) \leqslant u$ ($\forall u \in X^*$)，因此若 f 在 x 次可微，则 $D_+ f(x, \cdot) > -\infty$.

(iii) 若有 $r > f(x)$ 使 $(x,r) \in (\text{epi } f)^\circ$，则 f 在 x 处次可微.

(iv) 若 f 在 x 处连续，则 f 在 x 处次可微，且 $\partial f(x)$ 是弱*紧凸集.

证 (i)，(ii) 的证明是直接的(参考 1.5.3(i)).

(iii) 由 1.1.2 有 $0 \neq (v,\beta) \in X^* \times \mathbf{R}$，使得

$$v(x) + \beta f(x) \leqslant v(y) + \beta t \ (f(y) \leqslant t),$$

$v(x)+\beta f(x) < v(x)+\beta r$. 这推出 $\beta > 0, -v/\beta \in \partial f(x)$.

(iv) 由 f 在 x 处连续推出 $\forall r > f(x)$：$(x,r) \in (\text{epi } f)^\circ$，因此 $\partial f(x) \neq \varnothing$. 直接看出 $\partial f(x)$ 是弱*闭凸集. 由

$$-\Delta f(x, -z) \leqslant u(z) \leqslant \Delta f(x,z) \ (u \in \partial f(x))$$

及 f 在 x 连续得出 $\partial f(x)$ 等度连续，从而弱*紧(1.1.5). □

以下结果指明次微分与 G 导数的关系.

2.1.4 定理 设 f 凸且在 x 处连续，则 $\partial f(x) = \{u\} \Leftrightarrow u = \nabla f(x)$；当 $\dim X < \infty$ 时 $\partial f(x) = \{u\} \Leftrightarrow u = f'(x)$.

证 首先设 $\partial f(x) = \{u\}$. 取定 $z \in X\backslash\{0\}, \alpha \in [D_- f(x, z), D_+ f(x,z)]$. 定义 $u_\alpha: \mathbf{R}z \to \mathbf{R}, tz \to \alpha t$，则 $u_\alpha(y) \leqslant D_+ f(x, y)(\forall y \in \mathbf{R}z)$. 因 $D_+ f(x, \cdot)$ 是次线性的，故由 1.1.1 可设 u_α 是 X 上的线性泛函且 $u_\alpha \leqslant D_+ f(x, \cdot) \leqslant \Delta f(x, \cdot)$，于是由 f 在 x 连续得出 u_α 有界，从而 $u_\alpha \in \partial f(x)$(2.1.3). 因此有 $u_\alpha = u, \alpha = u(z) = D_\pm f(x,z)$，故 $u = \nabla f(x)$. 反之，若 $u = \nabla f(x)$，则用 2.1.3(ii)得出 $\partial f(x) = \{u\}$. 定理的后一结论由 1.5.3(i)推出. □

2.1.5 定理 设 $f, g: X \to \mathbf{R}$ 凸，f 在某点 $\bar{x} \in D_g$ 连续，则 $\partial(f+g) = \partial f + \partial g$.

证 取定 $x_0 \in D_f \cap D_g$，只需证

$$\partial(f + g)(x_0) \subset \partial f(x_0) + \partial g(x_0).$$

不妨设 $x_0 = 0, f(0) = g(0) = 0, 0 \in \partial(f + g)(0)$，待证 $0 \in$

$\partial f(0)+\partial g(0)$(一般情形由一适当代换得出). 在 $X\times\mathbf{R}$ 中对集 $A=\operatorname{epi}f$ 与 $B=\{(x,t): g(x)\leqslant -t\}$应用 1.1.2(i) 得 $\omega=(u,\beta)\in X^*\times\mathbf{R}$,使得$\langle\omega,A\rangle\leqslant\langle\omega,B\rangle$,$\langle\omega,A^\circ\rangle<\langle\omega,B\rangle$. 注意到 $(\hat{x},f(\hat{x})+1)\in A^\circ$,$(\hat{x},-g(\hat{x}))\in B$,易验证 $\beta<0$,$-u/\rho\in\partial f(0)$,$u/\rho\in\partial g(0)$,从而 $0\in\partial f(0)+\partial g(0)$. □

2.1.6 命题 设 $\bar{x}\in D$,则 $\bar{x}$ 是 f 的极小点 $\Leftrightarrow 0\in\partial f(\bar{x})\Leftrightarrow 0\in\partial f(\bar{x})+\partial\delta_D(\bar{x})\Leftrightarrow\partial f(\bar{x})\cap(D-\bar{x})^*\neq\varnothing$;若 D 是子空间,则 $\bar{x}$ 是 f 的极小点 $\Leftrightarrow\partial f(\bar{x})\cap D^\perp\neq\varnothing$.

证明是直接的(参考 2.1.2 之 4°). 2.1.6 推广了熟知的极值条件(参看 1.6.3(vi)). 所宜注意者,对于 $\bar{x}$ 为 f 的极小点,条件 $0\in\partial f(\bar{x})$不仅必要而且是充分的.

参考文献:[91,93,129,177,238].

§ 2.2 Clarke 次微分

本节设 $\Omega\subset X$ 为开集,$f:\Omega\to\mathbf{R}$,$\operatorname{Lip}f\leqslant K<\infty$.

2.2.1 定义 设 $x\in\Omega$. 称

$$f^\circ(x,z)\overset{\text{def}}{=}\limsup_{y\to x,t\downarrow 0}\frac{f(y+tz)-f(y)}{t}\tag{1}$$

为 f 在 x 的 **Clarke 广义方向导数**;称

$$\partial f(x)=\{u\in X^*: u\leqslant f^\circ(x,\cdot)\}\tag{2}$$

为 f 在点 x 处的 **Clarke 次微分**(或广义梯度). 若 $f^\circ(x,\cdot)=\mathrm{D}_+f(x,\cdot)$,则说 f 在 x 是正则的.

注 1° $f^\circ(x,z)$必存在且$|f^\circ(x,z)|\leqslant K|z|$. 用定义式(1)易验知 $f^\circ(x,\cdot)$是次线性的,且 $f^\circ(x,-z)=(-f)^\circ(x,z)$,因而 $f^\circ(x,\cdot)$为连续凸函数. 式(2)表明 $\partial f(x)$是 $f^\circ(x,z)$关于 z 在 $z=0$ 的次微分. 由 2.1.3(iv) 知 $\partial f(x)$是非空弱*紧凸集.

2° 若 f 凸，$x \in \Omega$，则 f 必在 x 正则：

$$D_+ f(x,z) \leqslant f^{\circ}(x,z) \leqslant \lim_{t \downarrow 0} \lim_{y \to x} \Delta f(y,tz) = D_+ f(x,z).$$

于是，对照(2)与 2.1.3(ii) 看出 $\partial f(x)$ 重合于次微分.

3° 若 x 是 f 的局部极小(或极大)点，则必定 $f^{\circ}(x,\cdot) \geqslant 0$，从而 $0 \in \partial f(x)$.

2.2.1 定义出两个映射 $f^{\circ}: \Omega \times X \to \mathbf{R}$ 与 $\partial f: \Omega \to 2^{X^*}$，它们的基本性质汇集于下.

2.2.2 定理 (i) f° 为 usc 且 Lip $f^{\circ}(x,\cdot) \leqslant K$，因而有 $R(\partial f) \subset \bar{B}_K(0) \subset X^*$.

(ii) ∂f 是 sw* 闭映射.

(iii) 若 $Q \subset X^*$ 为弱*紧凸集，$Q \neq \varnothing$，则 $\partial f(x) \subset Q \Leftrightarrow f^{\circ}(x,\cdot) \leqslant s_Q$，$\partial f(x) = Q \Leftrightarrow f^{\circ}(x,\cdot) = s_Q$，其中 s_Q 是集 Q 的**支承函数**，它定义为 $s_Q(x) = \sup \langle Q,x \rangle (x \in X)$.

(iv) 若 X 自反，则 ∂f 为 sw*-usc；若 $\dim X < \infty$，则 ∂f 为 usc.

证 (i)之证是直接的. 其次，由 $f^{\circ}(\cdot,z)$为 usc 推出(ii)(参考 1.3.1).

(iii) 令 $P = \partial f(x)$，则 $s_P \leqslant f^{\circ}(x,\cdot)$. $\forall z \in X$，依 $u(tz) = t f^{\circ}(x,z)$定义 $u: \mathbf{R}z \to \mathbf{R}$. 如同 2.1.4 之证，可设 $u \in X^*$ 且 $u \leqslant f^{\circ}(x,\cdot)$，于是 $u \in P$ 且 $u(z) = f^{\circ}(x,z)$. 故 $s_P = f^{\circ}(x,\cdot)$. 若 $P \subset Q$，则$f(x,\cdot) = s_P \leqslant s_Q$. 若 $P \not\subset Q$，则 $\exists u \in P \setminus Q$，于是由 1.1.3 有 $z \in X$，使 $\langle Q,z \rangle < u(z)$，从而 $s_Q(z) < u(z) \leqslant f^{\circ}(x,z)$. 这就证得 $P \subset Q \Leftrightarrow f^{\circ}(x,\cdot) \leqslant s_Q$. 类似地，可证明 $P = Q \Leftrightarrow f^{\circ}(x,\cdot) = s_Q$.

(iv) 若 ∂f 在某点 $x \in \Omega$ 非 sw*-usc(参看 1.3.1)，则 $\exists \varepsilon > 0, z \in X, x_n \to x, u_n \in \partial f(x_n)$，使得 $\forall u \in \partial f(x): |\langle u_n - u, z \rangle| \geqslant \varepsilon$. 因 $|u_n| \leqslant K$，不妨设 $u_n \rightharpoonup u$. 由(ii) 有 $u \in \partial f(x)$，这与 $|\langle u_n - u,$

$z\rangle| \geqslant \varepsilon$ 矛盾. □

2.2.2(iii)蕴涵等式

$$f^{\circ}(x;z) = \max\langle\partial f(x), z\rangle.$$

此式与式(2)一起表明,f° 与 ∂f 相互唯一确定.

若 $g: \Omega \to Y$ 在 x 处 G-可微,且对任何紧集 $C \subset X$,关于 $x \in C$ 一致地成立

$$\mathrm{D}g(x)z = \lim_{y \to x, t \downarrow 0} \frac{g(y+tz) - g(y)}{t}, \tag{3}$$

则说 g 在 x 处**严格可微**,称 $\mathrm{D}g(x)$ 为**严格导数**,记作 $\mathrm{D}_s g(x)$.

2.2.3 命题 设(3)对任何 $z \in X$ 成立,则 g 在 x 处严格可微 $\Leftrightarrow$ g 在 x 邻近为 Lip.

证 若 g 在 x 邻近不为 Lip,则有 $x_n, y_n \to x$, $|x_n - y_n| < n^{-1}$, $|g(x_n) - g(x_n)| > n|x_n - y_n|$. 令 $t_n = \sqrt{n}\,|x_n - y_n|$, $z_n = (y_n - x_n)/t_n$,则 $t_n \to 0$, $z_n \to 0$, $|\Delta g(x_n, t_n z_n)/t_n| > \sqrt{n}$. 这表明式(3)不对 $z = z_n$ 一致成立. 反之,若 g 在 x 非严格可微,则有 $z_n \to z$, $x_n \to x$, $t_n \to 0$,使得

$$|t_n^{-1}\Delta g(x_n, t_n z_n) - \mathrm{D}g(x)z_n| \geqslant \varepsilon > 0.$$

于是 g 在 x 邻近必不为 Lip,否则将有

$$\begin{aligned}\varepsilon \leqslant & |t_n^{-1}||g(x_n + t_n z_n) - g(x_n + t_n z)| \\ & + |t_n^{-1}\Delta g(x_n, t_n z) - \mathrm{D}g(x)z| \\ & + |\mathrm{D}g(x)||z - z_n| \to 0 \ (n \to \infty)!\end{aligned}$$ □

利用严格导数概念,可得与 2.1.4 类似的如下定理:

2.2.4 定理 若 $\nabla f(x)$ 存在,则 $\nabla f(x) \in \partial f(x)$. $\partial f(x) = \{u\} \Leftrightarrow u = \mathrm{D}_s f(x) \Leftrightarrow u = \nabla f(x)$ 且 f 在 x 正则. 若 $f \in C^1$,则 $\partial f(x) = f'(x) = \mathrm{D}_s f(x)$. 若 $\dim X < \infty$, ∂f 单值,则 $f \in C^1$.

证 若 $\partial f(x) = \{u\}$,则由 2.2.2(iii) 有 $f^{\circ}(x,\cdot) = u$. 而

$$\liminf_{y \to x, t \downarrow 0} \frac{1}{t}\Delta f(y, tz) = -f^{\circ}(x, -z) = u(z),$$

故由 2.2.3 有 $u = D_s f(x)$. 其余结论是明显的(参考1.3.1, 2.2.2). □

注 若 f 可微而非严格可微,则不必有 $\partial f(x) = f'(x)$. 例如,对 $f(x) = x^2\sin x^{-1}(x \neq 0)$, $f(0) = 0$,有 $\partial f(0) = [-1, 1]$,而 $f'(0) = 0$!

2.2.5 中值定理 若 $[x,y] \subset \Omega$,则

$$f(y) - f(x) \in \langle \partial f([x,y]), y - x\rangle. \tag{4}$$

证 令 $z_t = t'x + ty$, $g(t) = f(z_t)$,

$$\varphi(t) = g(t) + t[f(x) - f(y)],$$

则 $\varphi(0) = f(x) = \varphi(1)$. 因此 φ 有极值点 $t \in (0,1)$,从而

$$0 \in \partial\varphi(t) = \partial g(t) + f(x) - f(y).$$

于是只需证 $\partial g(t) \subset \langle \partial f(z_t), y - x\rangle \overset{\text{def}}{=} Q$, Q 必为紧区间. 由2.2.2 (iii),这归于证 $g^\circ(t,r) \leqslant \max rQ$ ($\forall r \in \mathbf{R}$),这由明显的不等式 $g^\circ(t,r) \leqslant f^\circ(z_t, r(y-x))$推出. □

2.2.6 推论 若 Ω 凸,则 f 凸 $\Leftrightarrow \partial f$ 单调.

证 显然 f 凸 $\Rightarrow \partial f$ 单调. 若 ∂f 单调, $x_1, x_2 \in \Omega$, $z = tx_1 + t'x_2$, $t \in J$,取 $y_i \in [x_i, z]$, $u_i \in \partial f(y_i)$,使

$$f(x_i) - f(z) = \langle u_i, x_i - z\rangle \quad (i = 1,2),$$

则

$$tf(x_1) + t'f(x_2) - f(z) = \beta\langle u_1 - u_2, y_1 - y_2\rangle \geqslant 0,$$

其中 $\beta = t|x_1 - z|/|y_1 - y_2| \geqslant 0$. 由此可见 f 凸. □

利用次微分表述极值条件,无疑是本书的重要课题. 前面已指明,若 $x \in \Omega$ 是 f 的极值点,则 $0 \in \partial f(x)$. 若 x 仅是 $f|D$ 的极小点, $D \subset \Omega$ 不必为开集,则条件"$0 \in \partial f(x)$"需作适当修改. 可考虑求助于指示函数 δ_D(参看 2.1.6 及下面的 2.2.9),但更精细的办法是使用**距离函数** $d_D(x) \overset{\text{def}}{=} d(x,D)$. 与 δ_D 相比, d_D 包含更多的信息且具有较好的性质,因而是更合用的工具. 容易直接验证: $d_D(x) = 0 \Leftrightarrow x \in \overline{D}$; D 凸 $\Rightarrow d_D$ 凸 $\Rightarrow \overline{D}$ 凸; $\mathrm{Lip}\,d_D \leqslant 1$; $R(\partial d_D) \subset B_1(0) \subset X^*$; $x \in D \Rightarrow 0 \in \partial d_D(x)$. 这些性质下面将不加指

明地反复使用.

2.2.7 定理 设 $D \subset \Omega$，$\bar{x}$ 是 $f|D$ 的极小点，$\varphi_\alpha = f + \alpha\, d_D$.

(i) 若 $\alpha \geqslant K$，则 $\bar{x}$ 是 φ_α 的极小点，且

$$0 \in \partial f(\bar{x}) + \alpha\, \partial d_D(\bar{x}). \tag{5}$$

(ii) 若 $\alpha > K$ 且 D 闭，则 $f|D$ 与 φ_α 有同样的极小点.

证 (i) 若 $\bar{x}$ 非 φ_α 之极小点，则 $\exists x \in \Omega, \varepsilon > 0$，使

$$\varphi_\alpha(x) < \varphi_\alpha(\bar{x}) - \alpha\varepsilon = f(\bar{x}) - \alpha\varepsilon.$$

取 $y \in D$，使 $|x-y| < d_D(x) + \varepsilon$，则

$$\begin{aligned} f(y) &\leqslant f(x) + K|x-y| \\ &\leqslant f(x) + \alpha\, d_D(x) + \alpha\varepsilon < f(\bar{x}), \end{aligned}$$

这与 $\bar{x}$ 是 $f|D$ 的极小点相矛盾. 因此 $\bar{x}$ 是 φ_α 的极小点. 于是，$0 \in \partial\varphi_\alpha(\bar{x})$，而这得出式(5).

(ii) 任取 φ_α 的极小点 x，今证 $x \in D$(从而 x 是 $f|D$ 的极小点). 令 $\beta = (\alpha + K)/2$，则 $K < \beta < \alpha$，于是

$$\varphi_\alpha(x) = \varphi_\alpha(\bar{x}) = \varphi_\beta(\bar{x}) \leqslant \varphi_\beta(x),$$

这推出 $\alpha\, d_D(x) \leqslant \beta d_D(x)$，因此 $d_D(x) = 0$，从而 $x \in D$. □

条件(5)可与 2.1.6 中的条件

$$0 \in \partial f(\bar{x}) + \partial\delta_D(\bar{x}) \tag{6}$$

相对比. 实际上，当 $\partial f(\bar{x})$ 是 Clarke 次微分时，(6)亦可用作 $f|D$ 的极值条件. 为说明这一点，首先给出定义 1.5.5 的非光滑推广如下：

2.2.8 定义 设 F, r, θ 如 1.5.5.

(i) 若 $\forall y \in D\backslash\{x\}, \forall u \in \partial f(x)$，有

$$F(y,u) + r\theta(x,y) \leqslant f(y) - f(x)\,[\, < f(y) - f(x)],$$

则说 f 在 x 关于 D 为广义(F,r,θ)-凸[广义(F,r,θ)-严格凸].

(ii) 若 $\forall y \in D, \exists u \in \partial f(x)$：

$$F(y,u) + r\theta(x,y) > 0 \Rightarrow f(y) > f(x),$$

则说 f 在 x 关于 D 为广义(F,r,θ)-拟凸.

(iii) 若 $\forall y \in D\backslash\{x\}, \exists u \in \partial f(x)$：$F(y,u) + r\theta(x,y) \geqslant 0$

$\Rightarrow f(y) \geqslant f(x)[f(y) > f(x)]$,则说 f 在 x 关于 D 为广义(F,r,θ)-伪凸[广义(F,r,θ)-严格伪凸].

f 在 D 上广义(F,r,θ)-凸等的意义自明.

若 $F(y,u) = \langle u,\eta(y)\rangle$,$\eta: D \to X$,则将 2.2.8 中的"$(F,r,\theta)$-"改为"$(\eta,r,\theta)$-";若 $r = 0$,则将"(F,r,θ)-"与"(η,r,θ)-"分别改为"(F,θ)-"与"(η,θ)-";若 $\theta(x,y) = |x-y|$,则在上述记号中省去 θ;若 $\eta(y) = y-x$,则在上述记号中省去 η. 为明确起见,分别给出 f 在 x 关于 D 为**广义凸、广义严格凸、广义拟凸、广义伪凸**与**广义严格伪凸**的直接刻画如下(对照§1.5(1)～(5)):

$$\forall y \in D: \langle \partial f(x), y - x\rangle \leqslant f(y) - f(x); \tag{7}$$

$$\forall y \in D\backslash\{x\}: \langle \partial f(x), y - x\rangle < f(y) - f(x); \tag{8}$$

$$\forall y \in D, \exists u \in \partial f(x): u(y - x) > 0 \Rightarrow f(y) > f(x); \tag{9}$$

$$\forall y \in D, \exists u \in \partial f(x): u(y - x) \geqslant 0 \Rightarrow f(y) \geqslant f(x); \tag{10}$$

$$\forall y \in D\backslash\{x\}, \exists u \in \partial f(x): u(y - x) \geqslant 0 \Rightarrow f(y) > f(x). \tag{11}$$

注意式(9)相当于 $\forall y \in D: f^\circ(x,y-x) > 0 \Rightarrow f(y) > f(x)$;式(10)与式(11)仿此. 以 $\eta(y)$ 代 $y-x$,可得广义 η-凸等概念的类似于式(7)～式(11)的刻画.

关于上述概念之间的关系,可以指出:若在§1.5(6)中每种凸性概念之前同时加上"广义",或"广义 η-",或"广义(F,r)-"等,则§1.5(6)仍然保持. 其次,若 $\partial f(x) = \nabla f(x)$,则 f 在 x 广义凸即 f 在 x 凸;余类推. 函数 $f(x) = \ln(1+|x|)$是广义严格伪凸的非凸函数的典型例子.

2.2.9 定理([107]) 设 $\bar{x} \in D \subset \Omega$,$D$ 凸. 若 $\bar{x}$ 是 $f|D$ 的极小点,则公式(6)成立(其中 $\partial\delta_D(\bar{x})$依 2.1.1);当 f 在 $\bar{x}$ 关于 D 为广义伪凸时其逆亦真.

证 设 $\bar{x}$ 是 $f|D$ 的极小点. 若式(6)不成立,则

$$\partial f(x) \cap (D - x)^* = \varnothing$$

(参考 2.1.2 之 4°).因 $\partial f(\bar{x})$ 弱*紧凸,而 $(D-\bar{x})^*$ 弱*闭凸,故由 1.1.3 有 $z \in X$:$\langle \partial f(\bar{x}), z\rangle < \langle (D-\bar{x})^*, z\rangle$,这推出 $f^\circ(\bar{x}, z) = \max\langle \partial f(x), z\rangle < 0$,$z \in X \cap (D-\bar{x})^{**} = \overline{\text{cone}}(D-\bar{x})$(1.4.3).取 $\alpha > 0$,$x \in D$,使 $\alpha(x-\bar{x})$ 充分接近 z,则

$$f^\circ(\bar{x}, \alpha(x - \bar{x})) = \alpha f^\circ(\bar{x}, x - \bar{x}) < 0.$$

另一方面,有

$$f^\circ(\bar{x}, x - \bar{x}) \geqslant \limsup_{t \downarrow 0} \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t} \geqslant 0$$

(此处用了 D 凸!),得出矛盾.可见式(6)成立.

反之,若式(6)成立,则存在 $u \in \partial f(\bar{x}) \cap (D-\bar{x})^*$,于是

$$f^\circ(\bar{x}, x - \bar{x}) \geqslant u(x - \bar{x}) \geqslant 0 \ (\forall\, x \in D).$$

因此,当 f 在 $\bar{x}$ 关于 D 为广义伪凸时 $f(D) \geqslant f(\bar{x})$,即 $\bar{x}$ 是 $f|D$ 的极小点. □

2.2.9 可看作 1.6.3(vi)的非光滑推广.

关于广义 (F, r)-凸性的应用参看 §4.6.

参考文献:[39,93,108,161,178,179,189,204,213,214].

§2.3 次微分规则

如同在经典微分学中一样,基本的"次微分规则"是链规则,其他许多规则由其导出.

2.3.1 定理 设 $\varphi: \mathbf{R}^m \to \mathbf{R}$,$g = (g_i): X \to \mathbf{R}^m$,$g$ 与 φ 分别在 $x \in X$ 与 $y = g(x)$ 邻近为 Lip,$f = \varphi \circ g$,则

$$\partial f(x) \subset \overline{\text{co}}\, \partial\varphi(y) \circ \partial g(x), \tag{1}$$

其中 $\overline{\text{co}}$ 记弱*闭凸包(下同),$\partial g(x) \stackrel{\text{def}}{=} \prod \partial g_i(x) \subset L(X, \mathbf{R}^m)$.以下每个条件推出式(1)为等式:(i) g_i ($1 \leqslant i \leqslant m$)在 x 及 φ 在 y 正则,$\partial\varphi(y) \subset \mathbf{R}^m_+$;(ii) $m = 1$,φ 在 y 严格可微;(iii) g 在 x 严格可

微，φ 在 y 正则. 在条件(i)或(iii)之下 f 在 x 正则；在条件(ii)或(iii)之下可去掉记号 $\overline{\mathrm{co}}$.

证 令 $Q = \partial\varphi(y) \circ \partial g(x)$，则 Q(从而 $\overline{\mathrm{co}}\, Q$)弱*紧. 只要证 $f^\circ(x, \cdot) \leqslant s_Q$(2.2.2). 取定 $z \in X$. $\forall \varepsilon > 0$，令

$$q_\varepsilon = \sup\left\{ \sum \alpha_i u_i(z) : (\alpha_i) \in \partial\varphi(B_\varepsilon(y)), u_i \in \partial g_i(B_\varepsilon(x)) \right\}, \tag{2}$$

$q = s_Q(z)$，只要证 $f^\circ(x,z) < q_\varepsilon + \varepsilon$，$q_\varepsilon \to q\ (\varepsilon \downarrow 0)$. 取 $x' \in B_\varepsilon(x)$，$t > 0$，使 $f^\circ(x,z) < t^{-1}\Delta f(x', tz) + \varepsilon$，$x' + tz \in B_\varepsilon(x)$，且 $g(x')$，$g(x' + tz) \in B_\varepsilon(y)$. 由 2.2.5 有

$$\begin{aligned}\Delta f(x', tz) &\in \langle \partial\varphi([g(x'), g(x' + tz)]), \Delta g(x', tz) \rangle \\ &\subset \langle \partial\varphi(B_\varepsilon(y)), \Delta g(x', tz) \rangle; \\ \Delta g_i(x', tz) &\in \langle \partial g_i(B_\varepsilon(x)), tz \rangle.\end{aligned}$$

由式(2)看出 $\Delta f(x', tz) \leqslant tq_\varepsilon$，从而 $f^\circ(x,z) < q_\varepsilon + \varepsilon$. 不妨设 $\mathrm{Lip}\ \varphi, \mathrm{Lip}\ g \leqslant K < \infty$. $\forall \delta > 0$，取 ε 充分小，使

$$\partial\varphi(B_\varepsilon(y)) \subset \partial\varphi(y) + B_\delta(0),$$

且

$$g_i^\circ(B_\varepsilon(x), \pm z) \leqslant g_i^\circ(x, \pm z) + K^{-1}\delta$$

(2.2.2)，则

$$\begin{aligned} q_\varepsilon &\leqslant \sup\left\{ \sum \sup\{\alpha_i u_i(z) : u_i \in \partial g_i(B_\varepsilon(x))\} : (\alpha_i) \in \partial\varphi(B_\varepsilon(y)) \right\} \\ &\leqslant \sup\left\{ \sum [g_i^\circ(x, \alpha_i z) + \delta] : (\alpha_i) \in \partial\varphi(B_\varepsilon(y)) \right\} \\ &\leqslant \sup\left\{ \sum \max\langle \partial g_i(x), \alpha_i z \rangle : (\alpha_i) \in \partial\varphi(y) + B_\delta(0) \right\} + m\delta \\ &\leqslant q + m\delta(K|z| + 1), \end{aligned}$$

这与 $q \leqslant q_\varepsilon$ 一起推出 $q_\varepsilon \to q\ (\varepsilon \downarrow 0)$. 于是式(1)成立.

若条件(i) 满足，则依上段记号有

$$\begin{aligned} q &= \max\left\{ \sum \max\langle \partial g_i(x), \alpha_i z \rangle : (\alpha_i) \in \partial\varphi(y) \right\} \\ &= \max\left\{ \sum \alpha_i g_i^\circ(x, z) : (\alpha_i) \in \partial\varphi(y) \right\} \end{aligned}$$

$$= \max\langle \partial\varphi(y), \omega\rangle \qquad (\omega = (\mathrm{D}_+ g_i(x,z)))$$

$$= \varphi^\circ(y,\omega) = \mathrm{D}_+ \varphi(y,\omega) = \mathrm{D}_+ f(x,z),$$

可见 $f^\circ(x,z) = q = \mathrm{D}_+ f(x,z)$，因而式(1)是等式，且 f 在 x 正则. 当条件(iii)满足时可类似处理. 若条件(ii)满足，则可设 $\alpha \overset{\mathrm{def}}{=} \mathrm{D}_s\varphi(y) \geqslant 0$，于是

$$q = \alpha g^\circ(x,z) = \limsup_{\xi \to x, t \downarrow 0} \frac{\alpha}{t} \Delta g(\xi, tz) = f^\circ(x,z),$$

这同样得出式(1)为等式. □

2.3.2 定理(Aubin-Clarke，1979) 设 $g: X \to Y, \varphi: Y \to \mathbf{R}$，$g$ 在 $x \in X$ 严格可微，φ 在 $y = g(x)$ 邻近为 Lip，$f = \varphi \circ g$，则

$$\partial f(x) \subset \partial\varphi(y) \circ \mathrm{D}g(x). \tag{3}$$

以下每个条件推出式(3)为等式：(i) φ 或 $-\varphi$ 在 y 正则；(ii) $\forall \varepsilon > 0, \exists \delta > 0: B_\delta(y) \subset \overline{gB_\varepsilon(x)}$；(iii) $R(\mathrm{D}g(x)) = Y$.

证 令 $A = \mathrm{D}g(x)$，$Q = A^* \partial\varphi(y)$，今只要证 $f^\circ(x,z) \leqslant s_Q(z)$. 这由以下推演得出：

$$f^\circ(x,z) = \limsup_{\xi \to x, t \downarrow 0} t^{-1}[\varphi(g(\xi + tz)) - \varphi(g(\xi))]$$

$$\leqslant \varphi^\circ(y, Az) + \limsup_{\xi \to x, t \downarrow 0} t^{-1}[\varphi(g(\xi + tz)) - \varphi(g(\xi) + tAz)]$$

$$\leqslant \varphi^\circ(y, Az) + \limsup_{\xi \to x, t \downarrow 0} \mathrm{const}\,|t^{-1}\Delta g(\xi, tz) - Az|$$

$$= \varphi^\circ(y, Az) = \max\langle \partial\varphi(y), Az\rangle = s_Q(z), \ z \in X.$$

若 φ 在 y 正则(当 $-\varphi$ 正则时仿此)，则

$$f^\circ(x,z) \leqslant \varphi^\circ(y, Az) = \mathrm{D}_+ \varphi(y, Az) = \mathrm{D}_+ f(x,z),$$

这得出 $f^\circ(x,z) = \mathrm{D}_+ f(x,z) = s_Q(z)(\forall z \in X)$，因此式(3)为等式，且 f 在 x 为正则. 若条件(ii)满足，则

$$\varphi^\circ(y, Az) = \limsup_{\eta \to y, t \downarrow 0} t^{-1}\Delta\varphi(\eta, tAz)$$

$$\leqslant \limsup_{\xi \to x, t \downarrow 0} t^{-1}\Delta\varphi(g(\xi), tAz) = f^\circ(x,z),$$

这同样得出 $f^\circ(x,z) = s_Q(z)$. 当 $R(A) = Y$ 时(ii)必满足. □

2.3.1 与 2.3.2 有一系列推论.

2.3.3 推论 设 $f, g_i: X \to \mathbf{R}(1 \leqslant i \leqslant m)$在 x 邻近为 Lip, $\alpha \in \mathbf{R}$,则

$$\partial(\alpha f)(x) = \alpha \partial f(x); \tag{4}$$

$$\partial\left(\sum g_i\right)(x) \subset \sum \partial g_i(x); \tag{5}$$

$$\partial(g_1 g_2)(x) \subset g_2(x)\partial g_1(x) + g_1(x)\partial g_2(x); \tag{6}$$

$$\partial(g_1/g_2)(x) \subset [g_2(x)\partial g_1(x) - g_1(x)\partial g_2(x)]/g_2^2(x). \tag{7}$$

当 $g_i(1 \leqslant i \leqslant m)$在 x 正则时式(5)是等式;当 g_i 在 x 正则且 $g_i(x) \geqslant 0(i = 1,2)$时式(6)是等式;式(7)要求 $g_2(x) \neq 0$;当 g_1 与 $-g_2$ 在 x 正则且 $g_1(x) \geqslant 0, g_2(x) > 0$ 时式(7)是等式.

以上结论可直接从 2.3.1 推出. 例如,令 $\varphi(y_1, y_2, \cdots, y_m) = \sum y_i, g = (g_i)$,对 $\sum g_i = \varphi \circ g$ 应用 2.3.1 即得式(5).

2.3.4 推论 设 $A \in L(X,Y), x \in X, \varphi: Y \to \mathbf{R}$ 在 Ax 邻近为 Lip 且在 Ax 正则,则 $\partial(\varphi \circ A)(x) = A^* \partial\varphi(Ax)$.

此由 2.3.2 推出. 注意连续凸(或凹)的 φ 满足 2.3.4 之条件.

2.3.5 推论 设 X 连续地并稠密地嵌入 $Y, x \in X, \varphi: Y \to \mathbf{R}$ 在点 x 邻近为 Lip, 则 $\partial(\varphi|X)(x) = \partial\varphi(x)$, 这意味着:每个 $u \in \partial(\varphi|X)(x)$可唯一地扩张为某个 $v \in Y^*$ 且 $v \in \partial\varphi(x)$.

这由应用 2.3.2 于 $\varphi|X = \varphi \circ g$ 得出,此处 $g: X \to Y$ 是嵌入映射.

2.3.6 定理 设 $\varphi: X = X_1 \times X_2 \to \mathbf{R}$ 在 $x = (x_1, x_2)$邻近为 Lip, $\partial_i\varphi(x)$记 $\varphi(x_1, x_2)$对 x_i 的广义梯度, $\pi_i: X^* \to X_i^*$ 为投影.

(i) $\partial_i\varphi(x) \subset \pi_i\partial\varphi(x), i = 1,2$.

(ii) 若 φ 在 x 正则,则 $\partial_i\varphi(x) = \pi_i\partial\varphi(x)(i=1,2)$,从而

$$\partial\varphi(x) \subset \partial_1\varphi(x) \times \partial_2\varphi(x).$$

证 (i) 令 $g(\cdot) = (\cdot, x_2)$,则 $\varphi(\cdot, x_2) = \varphi \circ g$,于是由 2.3.2得 $\partial_1\varphi(x) \subset \pi_1\partial\varphi(x)$. 同理 $\partial_2\varphi(x) \subset \pi_2\partial\varphi(x)$.

(ii) 设 $u = (u_1, u_2) \in \partial\varphi(x), u_i \in X_i^*(i = 1,2)$,今证 $u_1 \in$

$\partial_1\varphi(x)$(同理将有 $u_2 \in \partial_2\varphi(x)$). $\forall z \in X_1$,有

$$u_1(z) = u(z,0) \leqslant \varphi^\circ(x,(z,0))$$
$$= \mathrm{D}_+ \varphi(x,(z,0)) \leqslant \varphi_1^\circ(x,z)$$

(φ_1° 记 φ 对 x_1 的 Clarke 广义方向导数),可见 $u_1 \in \partial_1\varphi(x)$. □

上面建立的"次微分规则"无疑有助于次微分的计算. 然而,在一般情况下,次微分计算远不是容易的. 问题在于,$\partial f(x)$的定义(§2.2(2))不具备构造性. $f^\circ(x,z)$的定义(§2.2(1))虽然具有构造性,但§2.2(1)并不是一个易操作的算式. 不过,当 $X = \mathbf{R}^n$ 时能给出 $\partial f(x)$一个构造性表示.

2.3.7 定理 设 $f: \Omega \subset \mathbf{R}^n \to \mathbf{R}$ 为 C^{1-0} 函数,$S \subset \mathbf{R}^n$ 是(Lebesgue)零测集,N 记 f 的不可微点集. 则对每个 $x \in \Omega$ 有

$$\partial f(x) = \mathrm{co}\{\lim_k f'(x_k) : x_k \to x, x_k \overline{\in} S \cup N\}. \qquad (8)$$

约定"$f'(x_k)$收敛"这一限制已隐含于表达式中.

证 由一个著名定理(参考[32]),函数 f 几乎处处可微,因此 mes $N = 0$. 以 Q 记式(8)之右端,则易见 Q 是非空有界凸闭集,且 $Q \subset \partial f(x)$(2.2.4,2.2.2(ii)). 于是只要证 $f^\circ(x,\cdot) \leqslant s_Q$(2.2.2). 为此只需证:若 $\varepsilon > 0, 0 \neq z \in \mathbf{R}^n$,则

$$f^\circ(x,z) - \varepsilon \leqslant \limsup_{y \overline{\in} S\cup N, y \to x} \langle f'(y), z\rangle \overset{\text{def}}{=} \beta.$$

可设$|z| = 1$. 取 $\delta > 0$,使得 $\forall y \in B_{2\delta}(x) \backslash (S \cup N)$: $\langle f'(y), z\rangle < \beta + \varepsilon$. 由 Fubini 定理,对几乎所有的 $y \in B_\delta(x)$,$[y, y+\delta z] \cap (S \cup N)$是 1 维零测集;对这样的 y 与任何 $t \in (0,\delta)$,有

$$\frac{1}{t}\Delta f(y,tz) = \frac{1}{t}\int_0^t \langle f'(y+sz), z\rangle \mathrm{d}s \leqslant \beta + \varepsilon.$$

由连续性,不等式 $t^{-1}\Delta f(y,tz) \leqslant \beta + \varepsilon$ 可用于所有 $y \in B_\delta(x)$ 与 $t \in (0,\delta)$,因此 $f^\circ(x,z) \leqslant \beta + \varepsilon$. □

若令 $B_k = B_{1/k}(x) \backslash (S \cup N)$,则式(8)可表成

$$\partial f(x) = \mathrm{co} \bigcap_{k=1}^{\infty} \overline{f'(B_k)}. \qquad (9)$$

当 $f'(B_k)$不太复杂时,应用公式(9)计算 $\partial f(x)$是方便的.

2.3.8 例 1° 设 $f(x) = \max\{\min\{x_1, -x_2\}, x_2 - x_1\}$，$x = (x_1, x_2) \in \mathbf{R}^2$，即

$$f(x) = \begin{cases} x_1, & x_2 \leqslant 2x_1 \text{ 且 } x_2 \leqslant -x_1, \\ -x_2, & -2x_1 \leqslant 2x_2 \leqslant x_1, \\ x_2 - x_1, & x_1 \leqslant 2x_2 \text{ 或 } 2x_1 \leqslant x_2. \end{cases}$$

直接看出 $f'(x)$ 仅取值 $(1,0)$，$(0,-1)$ 与 $(-1,1)$，以此三点为顶点的实心三角形即 $\partial f(0)$. 其次，易知 $\partial_1 f(0) = [-1,0]$，$\partial_2 f(0) = [0,1]$，因而 $\partial f(0) \not\subset \partial_1 f(0) \times \partial_2 f(0)$，故 f 在 $x = 0$ 不是正则的(参考 2.3.6).

2° 设 $f(x) = x^2 \sin x^{-1} (x \neq 0)$，$f(0) = 0$. f 处处可微：

$$f'(x) = 2x\sin x^{-1} - \cos x^{-1} (x \neq 0),\ f'(0) = 0.$$

于是应用式(9)可得 $\partial f(0) = [-1,1]$.

今将 2.3.7 用于求 $\partial d_D(x)$. 首先引入概念：若 $D \subset \mathbf{R}^n$，$0 \neq z \in \mathbf{R}^n$，$z$ 可表为 $x' - x$，x 是 $\overline{D}$ 中唯一最接近 x' 的点，则说 z 在 x 垂直于 D，记作 $z \perp_x D$.

2.3.9 引理 设 $D \subset \mathbf{R}^n$，$x \in \mathbf{R}^n$，$d'_D(x) \neq 0$，则 $|d'_D(x)| = 1$，且 $\exists y \in \overline{D}$：$d'_D(x) = (x-y)/|x-y|$，$|x-y| d'_D(x) \perp_y D$.

证 取 $y \in \overline{D}$，使 $d_D(x) = |x-y|$，则

$$\begin{aligned} \langle d'_D(x), x - y \rangle &= \lim_{t \downarrow 0} t^{-1}[d_D(x) - d_D(x + t(y - x))] \\ &= |x - y|. \end{aligned} \tag{10}$$

必定 $x \neq y$（否则 $d'_D(x) = 0$），于是(10)与 $|d'_D(x)| \leqslant 1$ 一起推出 $|d'_D(x)| = 1$，$d'_D(x) = (x-y)/|x-y|$. 其余结论是明显的.

□

2.3.10 定理 设 $x \in \overline{D} \subset \mathbf{R}^n$，则 $\partial d_D(x) = \mathrm{co}(Q \cap \{0\})$，其中

$$Q = \{\lim z_k/|z_k| : z_k \perp_{x_k} D, x_k \to x, z_k \to 0\}.$$

证 由 2.3.7 与 2.3.9 有 $\partial d_D(x) \subset \mathrm{co}(Q \cup \{0\})$. 显然 $0 \in$

$\partial d_D(x)$，只要证 $Q \subset \partial d_D(x)$. 设 $z = \lim z_k/|z_k|$，$z_k \perp_{x_k} D$，$x_k \to x$，$z_k = y_k - x_k \to 0$，$d_D(y_k) = |y_k - x_k|$. 则 $|z_k| = d_D(y_k) - d_D(x_k) = \langle u_k, y_k - x_k \rangle$，$u_k \in \partial d_D(x'_k)$，$x'_k \in [x_k, y_k]$(2.2.5). 因 $|u_k| \leqslant 1$，故必 $u_k = z_k/|z_k| \to z$，$z \in \partial d_D(x)$(2.2.2). □

参考文献:[32,39,93,228,229].

§2.4 极大函数

对于由一有限运算得出的函数，如有限和、有限积及复合函数，上节给出了适当的"次微分规则". 对于由一"无限构成过程"作出的函数，如无限和、极限函数、上确界函数等，亦需类似的规则. 本节对所谓极大函数解决这一问题，所得结果对于导出非光滑最优性条件有重要作用.

设 T 是一紧度量空间，$\Omega \subset X$ 是一开集. 给定一函数族 $\{f_t \in C(\Omega) : t \in T\}$，考虑"极大函数" $f(x) = \max_T f_t(x)$. 为方便起见，也将 $f_t(x)$ 表为 $f(t,x)$. 假定 $f(\cdot, x)$ 为 usc 且 $\mathrm{Lip}\, f_t \leqslant K \leqslant \infty$，于是 $f: \Omega \to \mathbf{R}$ 有定义且 $\mathrm{Lip}\, f \leqslant K$. 约定

$$T_x = T(x) = \{t \in T : f_t(x) = f(x)\};$$

$$\partial_T f_t(x) = \overline{\mathrm{co}}\{u : u \text{ 是} \{u_n\} \text{的弱}^* \text{极限点},$$
$$u_n \in \partial f_{t_n}(x_n), t_n \to t, x_n \to x\}, \qquad (1)$$

其中 $\overline{\mathrm{co}}$ 记弱*闭凸包，本节中概如此.

2.4.1 定理(Clarke,1973) 任给 $x \in \Omega$，有

$$\partial f(x) \subset \overline{\mathrm{co}}\Big(\bigcup_{t \in T(x)} \partial_T f_t(x)\Big). \qquad (2)$$

若 $\forall t \in T_x$，f_t 在 x 正则，$\partial_T f_t(x) = \partial f_t(x)$，则 f 在 x 正则，且

$$\partial f(x) = \overline{\mathrm{co}}\Big(\bigcup_{t \in T(x)} \partial f_t(x)\Big)$$
$$= \left\{\int_T u_t \mathrm{d}\mu(t) : \mu \in P(T_x), u_t \in \partial f_t(x)\right\}, \qquad (3)$$

其中 $P(T_x)$ 记 T 上由 T_x 支承的概率 Radon 测度之全体，积分 $\int_T u_t \mathrm{d}\mu(t)$ 在弱* 收敛意义下存在.

证 首先证(2)式. 令 $Q = \bigcup_{t \in T(x)} \partial_T f_t(x)$，只需证 $f^\circ(x, \cdot) \leqslant s_Q$. 取定 $z \in X$. 取 $x_n \to x, \tau_n \downarrow 0$，使

$$\alpha_n \stackrel{\text{def}}{=} \tau_n^{-1} \Delta f(x_n, \tau_n z) \to f^\circ(x, z).$$

设 $t_n \in T(x_n + \tau_n z)$，则 $\alpha_n \leqslant \tau_n^{-1} \Delta f_{t_n}(x_n, \tau_n z) = u_n(z), u_n \in \partial f_{t_n}(y_n), y_n \in [x_n, x_n + \tau_n z]$(2.2.5). 可设 $t_n \to t \in T$, $\{u_n\}$ 有弱* 极限点 u. 因 $y_n \to x$，由式(1)有 $u \in \partial_T f_t(x)$. 由 $\alpha_n \leqslant u_n(z)$ 推出 $f^\circ(x, z) \leqslant u(z)$. 由

$$\begin{aligned} f_t(x) &\geqslant \limsup_n f_{t_n}(x) \\ &\geqslant \limsup_n [f(x_n + \tau_n z) - K|x_n + \tau_n z - x|] = f(x) \end{aligned}$$

得出 $f_t(x) = f(x)$，即 $t \in T_x$. 于是 $u \in Q$，从而有 $f^\circ(x, z) \leqslant s_Q(z)$.

其次在定理所给附加条件下证式(3). $\forall t \in T_x$，有

$$f_t'(x, z) = \lim_{\tau \downarrow 0} \frac{1}{\tau} \Delta f_t(x, \tau z) \leqslant \liminf_{\tau \downarrow 0} \frac{1}{\tau} \Delta f(x, \tau z).$$

于是

$$\begin{aligned} f^\circ(x, z) \leqslant s_Q(z) &= \sup_{t \in T(x)} \max \langle \partial f_t(x), z \rangle \\ &= \sup_{t \in T(x)} f_t'(x, z) \leqslant \liminf_{\tau \downarrow 0} \frac{1}{\tau} \Delta f(x, \tau z), \end{aligned}$$

这推出 f 在 x 正则. 由 Aubin 的一个结果可推知式(3)中第二个等号成立(参考[39, p. 89]). 在式(3)右端任取一元 $u = \int_T u_t \mathrm{d}\mu(t), u_t \in \partial f_t(x), \mu \in P(T_x)$，有

$$\begin{aligned} u(z) = \int_T u_t(z) \mathrm{d}\mu(t) &\leqslant \int_T f_t^\circ(x, z) \mathrm{d}\mu(t) \\ &= \int_T \lim_{\tau \downarrow 0} \Delta f_t(x, \tau z) \mathrm{d}\mu(t) \end{aligned}$$

$$= \lim_{\tau \downarrow 0} \frac{1}{\tau} \int_T \Delta f_t(x, \tau z) \mathrm{d}\mu(t)$$

$$\leqslant \lim_{\tau \downarrow 0} \frac{1}{\tau} \int_T \Delta f(x, \tau z) \mathrm{d}\mu(t) = f^\circ(x, z)$$

(其中用了控制收敛定理),这得出 $u \in \partial f(x)$. 因此式(3)成立. □

在 2.4.1 的应用中,"$\partial_T f_t(x) = \partial f_t(x)$"成立是重要的. 以下推论皆属于这种情况.

2.4.2 推论 若 $f(t, x)$对 t 连续、对 x 凸,则 $f(x)$凸,且式(3)成立.

证 直接看出 $f(x)$凸. 只需说明对 $t \in T_x$, u 是$\{u_n\}$的弱*极限点. 为记号简便,不妨设 $u_n \rightharpoonup u$. 于是

$$u(z) = \lim_n u_n(z) \leqslant \limsup_n \Delta f_{t_n}(x_n, z)$$

$$\leqslant \limsup_n [\Delta f_{t_n}(x, z) + 2K|x_n - x|] = \Delta f_t(x, z),$$

这表明 $u \in \partial f_t(z)$. □

2.4.3 推论 若 f_t 皆严格可微,且$(t, x) \to \mathrm{D}f_t(x)$连续,则 f 正则,且

$$\partial f(x) = \overline{\mathrm{co}}\{\mathrm{D}f_t(x) : t \in T_x\}$$

$$= \left\{\int_T \mathrm{D}f_t(x) \mathrm{d}\mu(t) : \mu \in P(T_x)\right\}, x \in \Omega. \qquad (4)$$

事实上,在所述条件下有 $\partial_T f_t(x) = \partial f_t(x) = \mathrm{D}f_t(x)$.

2.4.4 推论 设 $f_i \in C^{1-0}(\Omega)(1 \leqslant i \leqslant n)$, $f = \max_i f_i$,则

$$\partial f(x) \subset \mathrm{co}\{\bigcup_{i \in I(x)} \partial f_i(x)\}, \qquad (5)$$

其中 $I(x) = \{i : f_i(x) = f(x)\}$. 若 $\forall i \in I(x)$, f_i 在 x 正则,则 f 在 x 正则且式(5)为等式.

这是 2.4.1 的直接推论.

§2.5 切　　锥

在非光滑分析与最优化理论中，切锥概念有基本的重要性，其意义可与“光滑分析”中的切平面(或切线)概念相对比.

任给 $D \subset X$，$x \in D$，D 在 x 的**切锥**是某个锥，它在 x 邻近在某种意义上“接近于”D，因而也称为 D 在 x 处的**锥逼近**. 切锥的定义极为多样，见于文献者达几十种之多，但最常用者是如下几种：

$$V_D(x) = \{z: \text{当} t \downarrow 0, y \xrightarrow{D} x, w \to z \text{时} y + tw \in D\}; \quad (1)$$

$$F_D(x) = \{z: \text{当} t \downarrow 0, w \to z \text{时} x + tw \in D\}; \quad (2)$$

$$T_D(x) = \{z: \forall t_n \downarrow 0, x_n \xrightarrow{D} x, \exists z_n \to z, \text{使} x_n + t_n z_n \in D\} \quad (3)$$

$$K_D(x) = \{z: \exists t_n \downarrow 0, \exists z_n \to z, \text{使} x + t_n z_n \in D\}, \quad (4)$$

其中 $y \xrightarrow{D} x$ 表示 $y \in D$ 且 $y \to x$. 文献中，通常称 $V_D(x)$ 为**超切锥**，称 $F_D(x)$ 为**可行锥**(feasible cone)，称 $T_D(x)$ 为 **Clarke 切锥**，称 $K_D(x)$ 为 **Bouligand 切锥**. 若以 $A_D(x)$ 或 $A(D,x)$ 记锥式(1)～式(4)中的任何一个，则 $A_D(\cdot): D \to 2^X$ 是一“锥值”映射，它有性质：

$$A(D,x) = A(D - x, 0);$$

$$A(D,x) = A(D \cap B_\delta(x), x)(\forall \delta > 0);$$

$$x \in D \Rightarrow A(D,x) = X.$$

约定 $A_D^*(x) = (A_D(x))^*$，例如 $T_D^*(x) = (T_D(x))^*$. 称 $N_D(x) \overset{\text{def}}{=} -T_D^*(x)$ 为 D 在 x 的**法锥**.

不难直接验证：$V_D(x)$ 与 $F_D(x)$ 是开锥，$T_D(x)$ 与 $K_D(x)$ 是闭锥且 $T_D(x)$ 是凸锥；

$$V_D(x) \subset F_D(x) \cap T_D(x) \subset F_D(x) \cup T_D(x) \subset K_D(x). \quad (5)$$

(5)中的包含都可能是真包含. 若 $T_D(x) = K_D(x)$，则说 D 在 x

是正则的.

2.5.1 例 设 $f(x) = x\sin x^{-1}(x \neq 0)$, $f(0) = 0$, 令 $D = \text{epi}\, f \subset \mathbf{R}^2$. 由式(1)～式(4)直接看出：$V_D(0) = \varnothing$;

$$F_D(0) = \{x \in \mathbf{R}^2 : |x_1| < x_2\};$$
$$T_D(0) = \{0\} \times \mathbf{R}_+;$$
$$K_D(0) = \{x \in \mathbf{R}^2 : |x_1| \geqslant -x_2\}.$$

切锥的一些简单性质汇集于下：

2.5.2 命题 设 $(x,y) \in D\times E \subset X\times Y, D_i \subset X$.

(i) $F(\bigcap_1^n D_i, x) = \bigcap_1^n F(D_i, x)$; $F(D_1, x) \bigcap K(D_2, x) \subset K(D_1 \bigcap D_2, x)$; $V(D_1, x) \bigcap T(D_2, x) \subset T(D_1 \bigcap D_2, x)$.

(ii) 对 $A = V, F, T, N$ 有 $A_{D\times E}(x,y) = A_D(x)\times A_E(y)$；若 D 在 x(或 E 在 y)正则，则 $K_{D\times E}(x,y) = K_D(x)\times K_E(y)$.

(iii) $F_D(x) \subset \text{cone}(D^\circ - x)$；当 D 凸时等式成立；当 D 凸且 $D^\circ \neq \varnothing$ 时 $F_D^*(x) = (D-x)^*$.

(iv) 若 D 是凸锥，则 $D \subset T_D(0)$；若 D 是闭锥，则 $K_D(0) \subset D$；若 D 是闭凸锥，则 $T_D(0) = K_D(0) = D$.

证 只证(iii)，其余是明显的. 易见 $F_D(x) \subset \text{cone}(D^\circ - x)$. 若 D 凸，$z = \tau(y-x), \tau > 0, y \in D^\circ$，则当 $t \downarrow 0, w \to z$ 时有

$$x + tw = (1-\tau t)x + \tau t[y + \tau^{-1}(w-z)] \in D,$$

可见 $z \in F_D(x)$. 故得 $F_D(x) = \text{cone}(D^\circ - x)$. 若再设 $D^\circ \neq \varnothing$，则 $F_D^*(x) = (D^\circ - x)^* = (D-x)^*$. □

给定 $f: X \to \mathbf{R}$，"水平集"$D = \{x : f(x) \leqslant f(\bar{x})\}$在 $\bar{x}$ 的切锥可能对 f 在 $\bar{x}$ 邻近的局部性质提供重要信息，因此特别值得考察. 首先给出

2.5.3 引理 设 f 在 $\bar{x}$ 邻近为 Lip, $0 \overline{\in} \partial f(\bar{x})$, $K = \{z : f^\circ(\bar{x}, z) < 0\}$. 则 $K \neq \varnothing, \overline{K} = \{z : f^\circ(\bar{x}, z) \leqslant 0\}, K^* = (\overline{K})^* = \mathbf{R}_-\partial f(\bar{x})$，且

$$\begin{cases} K \subset V_D(\bar{x}) \subset F_D(x), \\ \bar{K} \subset T_D(\bar{x}) \subset K_D(\bar{x}); \end{cases} \tag{6}$$

$$\begin{cases} \mathbf{R}_- \partial f(\bar{x}) \supset V_D^*(\bar{x}) \supset F_D^*(\bar{x}), \\ \mathbf{R}_- \partial f(\bar{x}) \supset T_D^*(\bar{x}) \supset K_D^*(\bar{x}). \end{cases} \tag{7}$$

若 f 在 $\bar{x}$ 正则,则以上包含皆为等式.

以上结果不难直接证明.不过此处略去直接证明,因它可作为 4.1.8 的推论得出(参看 4.1.9).

若令 $\varphi(x,r) = f(x) - r$,则 f 为 Lip 推出 φ 为 Lip;

$$\varphi^\circ((x,r),(z,s)) = f^\circ(x,z) - s,$$

$$\partial\varphi(x,f(x)) = \partial f(x) \times \{-1\};$$

f 在 x 正则推出 φ 在 $(x,f(x))$ 正则;

$$\text{epi } f = \{(x,r) : \varphi(x,r) \leqslant 0\}.$$

以上结论结合 2.5.3 得出:

2.5.4 定理 设 f 在 $\bar{x}$ 邻近为 Lip,则

$$T_{\text{epi } f}(\bar{x},f(\bar{x})) = \text{epi } f^\circ(\bar{x},\cdot); \tag{8}$$

$$N_{\text{epi} f}(\bar{x},f(\bar{x})) = \mathbf{R}_+ (\partial f(\bar{x}) \times \{-1\}); \tag{9}$$

$$\partial f(\bar{x}) = \{u : (u,-1) \in N_{\text{epi} f}(\bar{x},f(\bar{x}))\}. \tag{10}$$

若 f 在 $\bar{x}$ 正则,则 epi f 在 $(\bar{x},f(\bar{x}))$ 正则.

证 由 2.5.3 有

$$\begin{aligned} \text{epi } f^\circ(\bar{x},\cdot) &= \{(z,s) : f^\circ(\bar{x},z) \leqslant s\} \\ &= \{(z,s) : \varphi^\circ((\bar{x},f(\bar{x})),(z,s)) \leqslant 0\} \\ &\subset T_{\text{epi } f}(\bar{x},f(\bar{x})). \end{aligned}$$

若 $(z,s) \in T_{\text{epi } f}(\bar{x},f(\bar{x}))$, $x_n \to \bar{x}$, $t_n \downarrow 0$,则 $\exists (z_n,s_n) \to (z,s)$: $(x_n,f(x_n)) + t_n(z_n,s_n) \in \text{epi } f$,即 $f(x_n + t_n z_n) \leqslant f(x_n) + t_n s_n$. 于是

$$\limsup_n t_n^{-1}[f(x_n + t_n z) - f(x_n)]$$

$$\leqslant \limsup_n t_n^{-1}[f(x_n + t_n z) - f(x_n + t_n z_n) + t_n s_n] = s,$$

这推出 $f^\circ(\bar{x},z)\leqslant s$，即 $(z,s)\in \text{epi}\ f^\circ(\bar{x},\cdot)$. 因此式(8)成立. 余下的结论是明显的. □

2.5.5 例 设 $u\in X^*$, $A\in L(X,Y)$, $f(x)=u(x)+|Ax|$. 则 $\partial f(0)=u+A^*B$，此处 $B\subset Y^*$ 是闭单位球(参考 2.1.2, 2.1.5, 2.3.4)；

$$f^\circ(0,z)=\max\langle u+A^*B,z\rangle=u(z)+\max\langle B,Az\rangle$$
$$=u(z)+|Az|=f(z).$$

于是由 2.5.4 有 $T_{\text{epi}f}(0,0)=\text{epi}\ f$.

2.5.6 定理 设 $\bar{x}\in D\subset X$，则

$$T_D(\bar{x})=\{z: d_D^\circ(\bar{x},z)=0\}$$
$$\subset\{z: d_D(\bar{x}+tz)=o(t)(t\downarrow 0)\};\tag{11}$$
$$N_D(\bar{x})\supset \mathbf{R}_+\,\partial d_D(\bar{x})\supset \partial\delta_D(\bar{x});\tag{12}$$
$$\partial d_D(\bar{x})\subset N_D(\bar{x})\cap \bar{B}_1(0).\tag{13}$$

若 D 凸，则(11)～(13)中的包含是等式.

证 首先设 $z\in T_D(x)$. 若 $x_n\to\bar{x}$, $t_n\downarrow 0$，取 $y_n\in D$，使 $|x_n-y_n|=d_D(x_n)+o(t_n)$，则 $y_n\to\bar{x}$. 于是 $\exists z_n\to z$: $y_n+t_nz_n\in D$，从而

$$\limsup_n t_n^{-1}[d_D(x_n+t_nz)-d_D(x_n)]$$
$$\leqslant\limsup_n t_n^{-1}[|x_n+t_nz-y_n-t_nz_n|-d_D(x_n)]\leqslant 0.$$

这推出 $d_D^\circ(\bar{x},z)=0$，更有 $d_D(\bar{x}+tz)=o(t)(t\downarrow 0)$. 反之，假设 $d_D^\circ(\bar{x},z)=0$, $x_n\xrightarrow{D}\bar{x}$, $t_n\downarrow 0$，则

$$\lim_n t_n^{-1}d_D(x_n+t_nz)=\lim_n t_n^{-1}[d_D(x_n+t_nz)-d_D(x_n)]=0,$$

即 $d_D(x_n+t_nz)=o(t_n)(n\to\infty)$. 取 $y_n\in D$，使得 $|x_n+t_nz-y_n|=o(t_n)$，则 $z_n\stackrel{\text{def}}{=}t_n^{-1}(y_n-x_n)\to z$, $x_n+t_nz_n=y_n\in D$，可见 $z\in T_D(\bar{x})$. 于是(11)得证.

由式(11)推出 $\langle\partial d_D(\bar{x}),T_D(\bar{x})\rangle\leqslant 0$，这就有 $\mathbf{R}_+\partial d_D(\bar{x})\subset N_D(\bar{x})$. 若 $u\in\partial\delta_D(\bar{x})=(\bar{x}-D)^*$，则 $\bar{x}$ 是 $-u|D$ 的极小点，于

是 $0 \in -u+|u|\partial d_D(x)$(2.2.7)，从而 $u \in \mathbf{R}_+\partial d_D(\bar{x})$. 这证得式(12). 由式(12)与 Lip $d_D \leqslant 1$ 得出式(13).

下面设 D 凸，从而 d_D 凸. 因 d_D 在 $\bar{x}$ 正则，故(11)为等式. 任给 $u \in N_D(\bar{x})$，$y \in D$. 若 $t_n \downarrow 0$，$x_n \xrightarrow{D} \bar{x}$，则 $y - x_n \to y - \bar{x}$；由 D 凸有 $x_n + t_n(y - x_n) \in D$，因此 $y - \bar{x} \in T_D(\bar{x})$，于是 $N_D(\bar{x}) \subset (\bar{x} - D)^* = \partial\delta_D(\bar{x})$，故(12)为等式. 若 $u \in N_D(\bar{x}) \cap \overline{B}_1(0)$，则有 $u(D) \leqslant u(\bar{x})$. 于是

$$
\begin{aligned}
u(z) &\leqslant \inf_{x \in D}[u(x) + u(z - x)] \\
&\leqslant u(\bar{x}) + d_D(z) \quad (\forall z \in X),
\end{aligned}
$$

可见 $u \in \partial d_D(\bar{x})$. 故式(13)为等式. □

结合式(12)与 2.2.7 得出：

2.5.7 定理 设 $f: \Omega \subset X \to \mathbf{R}$ 为 Lip，$D \subset \Omega$，$\bar{x}$ 是 $f|D$ 的极小点，则 $\partial f(\bar{x}) \cap T_D^*(\bar{x}) \neq \varnothing$.

给出一个直接证明如下. 若 $\partial f(\bar{x}) \cap T_D^*(\bar{x}) = \varnothing$，则由 1.1.3 有 $z \in X$：$\langle \partial f(\bar{x}), z\rangle < \langle T_D^*(\bar{x}), z\rangle$. 这推出 $\langle \partial f(\bar{x}), z\rangle < 0$，从而 $f^\circ(\bar{x}, z) < 0$；$z \in T_D(\bar{x})$(注意 $T_D(\bar{x})$是闭凸锥！). 任取 $t_n \downarrow 0$，由 $z \in T_D(\bar{x})$有 $z_n \to z$，使得 $x + t_n z_n \in D$，从而有 $f(\bar{x} + t_n z_n) \geqslant f(\bar{x})$. 于是

$$
\begin{aligned}
f^\circ(\bar{x}, z) &\geqslant \limsup_n \frac{f(\bar{x} + t_n z) - f(\bar{x})}{t_n} \\
&= \limsup_n \frac{f(\bar{x} + t_n z_n) - f(\bar{x})}{t_n} \geqslant 0,
\end{aligned}
$$

得出矛盾. 依以上证法，2.5.7 中只需 f 为局部 Lip，$\bar{x}$ 是 $f|D$ 的局部极小点.

最后指出，在 $V_D(x)$与 $T_D(x)$之间，存在由下述定理表达的准确关系.

2.5.8 定理(Rockafellar，1980) 若 $V_D(x) \neq \varnothing$，则 $V_D(x) = (T_D(x))^\circ$；当 $D \subset \mathbf{R}^n$ 闭时不必假定 $V_D(x) \neq \varnothing$.

证 显然$V_D(x) \subset (T_D(x))^{\circ}$. $\forall y \in (T_D(x))^{\circ}$,取定一个$z \in V_D(x)$($V_D(x) \neq \varnothing$用于此!),当$t \downarrow 0$时有

$$y = (y - tz) + tz \in T_D(x) + V_D(x).$$

今证$T_D(x) + V_D(x) \subset V_D(x)$(由此可得出$(T_D(x))^{\circ} \subset V_D(x)$). 任取$y \in T_D(x)$, $z \in V_D(x)$. 设$t_n \downarrow 0$, $x_n \xrightarrow{D} x$, $w_n \to y + z$. 由$y \in T_D(x)$有$y_n \to y$,使$x_n + t_n y_n \in D$. 由$x_n + t_n y_n \xrightarrow{D} x$, $w_n - y_n \to z \in V_D(x)$得出,当$n \to \infty$时$(x_n + t_n y_n) + t_n(w_n - y_n) = x_n + t_n w_n \in D$,可见$y + z \in V_D(x)$.

若$D \subset \mathbf{R}^n$闭,且$z \in (T_D(x))^{\circ}$,则$\langle N_D(x)\backslash\{0\}, z\rangle < 0$ (1.4.4). 结合式(12)得$\max\langle S^{n-1} \cap \partial d_D(x), z\rangle < 0$. 于是,当$y$邻近$x$且$d'_D(y)$存在、非零时$\langle d'_D(y), z\rangle < 0$(2.3.7, 2.3.9). 基于此,用一个类似于2.3.7之证的论证得出:对几乎所有邻近x的$y \in D$,当$t \downarrow 0$, $w \to z$时有

$$d_D(y + tw) = \int_0^t \langle d'_D(y + sw), w\rangle \mathrm{d}s \leqslant 0,$$

即$y + tw \in D$. 由D闭,这得出$z \in V_D(x)$. □

参考文献:[15, 20, 39, 44, 60, 93, 160, 162, 178, 217～219, 221, 226, 227, 230].

第三章　择一定理

如下形式的断言称为**择一定理**：命题 P 与命题 Q 恰有一个成立（或者说：命题 P 与 Q 两择一）。以上断言相当于："P 的否定" $\Leftrightarrow Q$. 因此，任何等价性断言都可表为择一定理. 在这种一般的意义上，择一定理遍及整个数学. 不过，最优化理论中所用的择一定理通常涉及某个不等式系统的可解性，因此亦称为**可解性定理**. 这类定理被有效地应用于导出最优性条件与对偶性结果，因而成为最优化理论中的基本工具，且形成了一个被广泛注意的独立研究领域.

本章中，X, Y, Z, W 记给定的 LCS（当要求为 B- 空间时将予指明），假定 W, Y 中已分别由闭凸锥 W_+, Y_+ 导入序 $\leqslant$. 其次，设

$$F: X \to 2^W, \quad G: X \to 2^Y, \quad H: X \to 2^Z;$$

$$f: D \to W, \quad g: D \to Y, \quad h: D \to Z, D \subset X;$$

$$T \in L(X, W), A \in L(X, Y), B \in L(X, Z).$$

为证明命题 P 与 Q 两择一，只需指明：P 与 Q 互斥，且 P 不真 $\Rightarrow Q$ 为真. 对于本章所涉及的择一定理，其互斥性多半显而易见，因而其证明概予省略.

§3.1　Farkas 引理

G 的**对偶映射** $G^*: Y^* \to 2^{X^*}$ 定义如下（参考[18]）：

$$G^*\lambda = \{u \in X^*: (-u, \lambda) \in (\mathrm{Gr}\, G)^*\}, \lambda \in Y^*. \quad (1)$$

不难验证，对于 $A \in L(X, Y)$，A^* 就是通常的对偶算子.

给定凸锥 $K \subset Y$，记 $K_- = -K$. 本节的中心课题是，考察在

什么条件下成立等式

$$-(G^{-1}K_-)^* = G^*K^*. \tag{2}$$

式(2)相当于对任给 $f\in X^*$,以下命题两择一:(P) $\exists x\in D, y\in Gx$: $f(x)>0, y\leqslant 0$(此处序 $\leqslant$ 由锥 K 导入);(Q) $\exists\lambda\in K^*, \forall(x,y)\in \mathrm{Gr}\,G$: $f(x)\leqslant\lambda(y)$. 可见,建立等式(2)是一个择一定理问题. 当 $\mathrm{Gr}\,G$ 是 $X\times Y$ 的子空间时称 G 为线性映射(这等价于:$Gx+Gy\subset G(x+y), \alpha Gx=G(\alpha x), 0\in G(0), \forall x,y\in X, \alpha\neq 0$). 若 G 是线性的,则式(2)之左端 $=(G^{-1}K)^*$,因而式(2)可写成更整齐的形式:$(G^{-1}K)^*=G^*K^*$.

本节的基本结果如下.

3.1.1 定理 以下每个条件蕴涵等式(2):

(i) G 是 K-凸映射,$G(\alpha x)=\alpha Gx(\alpha>0)$,$\mathrm{epi}\,G$ 是闭集且 G^*K^* 弱*闭;

(ii) G 是线性映射且 $K^\circ\cap R(G)\neq\emptyset$;

(iii) K 与 $\mathrm{Gr}\,G$ 皆为闭凸锥且 $(\mathrm{Gr}\,G)^*+0\times K^*$ 弱*闭.

证 首先指出:$-(G^{-1}K_-)^*\supset G^*K^*$ 无条件成立. 事实上,若 $\lambda\in K^*, u\in G^*\lambda$,则 $(-u,\lambda)\in(\mathrm{Gr}\,G)^*$,于是对任给 $y\in K_-\cap Gx$(这意味着 $x\in G^{-1}K_-$),有 $0\leqslant -u(x)+\lambda(y)\leqslant -u(x)$,从而 $u\in-(G^{-1}K_-)^*$.

下面假定 $u\in-(G^{-1}K_-)^*$,要证 $u\in G^*K^*$.

首先设条件(i)满足. 若 $u\notin G^*K^*$,则由 G^*K^* 弱*闭有 $x\in X$: $u(x)>\langle G^*K^*,x\rangle$,这推出 $u(x)>0$, $\langle G^*K^*,x\rangle\leqslant 0$. 易见 $\mathrm{epi}\,G=\mathrm{Gr}\,G+0\times K$,于是

$$\begin{aligned}(\mathrm{epi}\,G)^* &= (\mathrm{Gr}\,G)^*\cap(0\times K)^* \\ &= (\mathrm{Gr}\,G)^*\cap(X^*\times K^*)\end{aligned}$$

(1.4.6). 任给 $(v,\lambda)\in(\mathrm{epi}\,G)^*$,必有 $\lambda\in K^*$, $(v,\lambda)\in(\mathrm{Gr}\,G)^*$;从而 $-v\in G^*\lambda\subset G^*K^*$, $-v(x)\leqslant 0$. 于是 $\langle(v,\lambda),(x,0)\rangle=v(x)\geqslant 0$. 这表明 $(x,0)\in(\mathrm{epi}\,G)^{**}$. 因在条件(i)下 $\mathrm{epi}\,G$ 是闭凸锥,故用 1.4.3 得 $(x,0)\in\mathrm{epi}\,G$. 这推出 $0\in Gx+$

K, $x \in G^{-1}K_-$. 但 $u(x) > 0$, 这与 $-u \in (G^{-1}K_-)^*$ 相矛盾.

其次设条件(ii)满足. 由 $u \in (G^{-1}K)^*$ 推出 $(u,0) \in (\mathrm{Gr}\, G \cap (X \times K))^* \stackrel{\mathrm{def}}{=} \Phi$. 由 $K^\circ \cap R(G) \neq \varnothing$ 推出 $\mathrm{Gr}\, G \cap (X \times K^\circ) \neq \varnothing$, 于是在 $X \times Y$ 中对子空间 $\mathrm{Gr}\, G$ 与锥 $X \times K$ 应用1.4.7得 $\Phi = (\mathrm{Gr}\, G)^* + (X \times K)^* = (\mathrm{Gr}\, G)^\perp + 0 \times K^*$. 因此有 $\lambda \in K^*$, 使得 $(u,0) - (0,\lambda) = (u, -\lambda) \in (\mathrm{Gr}\, G)^\perp$, 从而 $u \in G^*\lambda \subset G^*K^*$.

最后设条件(iii)满足, 则对闭凸锥 $\mathrm{Gr}\, G$ 与 $X \times K_-$ 应用1.4.7得 $\Phi \stackrel{\mathrm{def}}{=} (\mathrm{Gr}\, G \cap (X \times K_-))^* = (\mathrm{Gr}\, G)^* - 0 \times K^*$. 由 $u \in -(G^{-1}K_-)^*$ 推出 $(-u,0) \in \Phi$, 于是存在 $\lambda \in K^*$, 使 $(-u,\lambda) \in (\mathrm{Gr}\, G)^*$, 因此 $u \in G^*\lambda \subset G^*K^*$. □

关于3.1.1可参考[107].

若以 $A \in L(X,Y)$ 代 G, 则(2)式成为

$$(A^{-1}K)^* = A^*K^*. \tag{3}$$

3.1.2 定理 以下每个条件蕴涵等式(3):

(i) K 是闭凸锥且 A^*K^* 弱*闭;

(ii) $K^\circ \cap R(A) \neq \varnothing$;

(iii) K 是闭凸锥且 $(\mathrm{Gr}\, A)^\perp + 0 \times K^*$ 弱*闭;

(iv) K 是闭凸锥且 $K + R(A) = Y$[128, Th. 2.1].

证 当条件(i)～(iii)满足时, 结果直接由3.1.1得出. 余下只要证条件(iv) $\Rightarrow A^*K^*$ 弱*闭. 设有网 $\{\lambda_\sigma\} \subset K^*$, 使 $A^*\lambda_\sigma \xrightarrow{w^*} u$. 由 $K + R(A) = Y$ 知, $\forall y \in Y$, $\exists y_i \in K$, $x_i \in X (i = 1,2)$: $y = y_1 + Ax_1 = -y_2 + Ax_2$, 于是 $\langle A^*\lambda_\sigma, x_1\rangle = \lambda_\sigma(y - y_1) \leqslant \lambda_\sigma(y) \leqslant \lambda_\sigma(y + y_2) = \langle A^*\lambda_\sigma, x_2\rangle$. 这表明 $\{\lambda_\sigma\}$ 弱有界. 由1.1.5, 有子网 $\lambda_{\sigma'} \xrightarrow{w^*} \lambda \in K^*$, 从而 $u = A^*\lambda \in A^*K^*$. 因此 A^*K^* 弱*闭. □

若取 $A \in L(\mathbf{R}^n, \mathbf{R}^m)$, $K = \mathbf{R}^m_+$, 则 $K^* = \mathbf{R}^m_+$, $A^* = A^{\mathrm{T}} \stackrel{\mathrm{def}}{=} M$. 今指明 $M\mathbf{R}^m_+$ 闭. 可设 $R(M) = \mathbf{R}^n$ (否则以 $R(M)$ 代 $\mathbf{R}^n$). 设 $\{x_k\} \subset \mathbf{R}^m_+$, $Mx_k \to y$. 取 $x \in M^{-1}y$. 若 $x \notin \mathbf{R}^m_+$, 则 $\exists z_k \in [x, x_k] \cap$

$\partial\mathbf{R}^m_+$, $Mz_k\in[y,Mx_k]$，于是 $Mz_k\to y$. 用一归纳假设得出 $M(\partial\mathbf{R}^m_+)$ 闭，因此 $y\in M(\partial\mathbf{R}^m_+)\subset M\mathbf{R}^m_+$，即 $M\mathbf{R}^m_+$ 是闭集. 于是由 3.1.2 得出：

3.1.3 推论 若 $A\in L(\mathbf{R}^n,\mathbf{R}^m)$，则 $(A^{-1}\mathbf{R}^m_-)^*=A^{\mathrm{T}}\mathbf{R}^m_+$.

3.1.3 就是经典的 Farkas 引理，而 3.1.1 与 3.1.2 则可看作 Farkas 引理的推广.

3.1.4 推论 设 g 是凸映射且为 $*$ lsc（这意味着 $\forall\lambda\in Y^*_+$：$\lambda\circ g$ 为 lsc），$g(\alpha x)=\alpha g(x)(\alpha>0)$，则

$$-(g^{-1}Y_-)^*=\{u\in X^*:(u-\lambda g)|D\leqslant 0(\forall\lambda\in Y^*_+)\}$$
$$=\bigcup_{\lambda\in Y^*_+}\partial(\lambda g)(0)\qquad(\text{若 }D=X),$$

只要等式右端为弱 * 闭集.

证 用 3.1.1，只需说明 epi g 为闭集. $\forall\lambda\in Y^*_+$，λg 凸且为 lsc，因而为 wlsc(1.5.2). 由此可推出 epi g 弱闭；再由 1.1.4 得出 epi g 为闭集. □

3.1.4 的一个类似的形式见[107].

结合 2.2.9 与 3.1.1 ~ 3.1.3 得出：

3.1.5 推论 设 $f:X\to\mathbf{R}$ 在 $x=0$ 对某集 E 为广义伪凸.

(i) 若 3.1.1 的条件 (i) ~ (iii) 之一满足，$G^{-1}K_-$ 凸且 $G^{-1}K_-\subset E$，则 $f(G^{-1}K_-)\geqslant f(0)\Leftrightarrow 0\in\partial f(0)+G^*K^*$.

(ii) 若 3.1.2 的条件 (i) ~ (iv) 之一满足且 $A^{-1}K_-\subset E$，则 $f(A^{-1}K_-)\geqslant f(0)\Leftrightarrow 0\in\partial f(0)+A^*K^*$.

(iii) 若 3.1.3 的条件满足且 $g^{-1}Y_-\subset E$，$D=X$，则 $f(g^{-1}Y_-)\geqslant f(0)\Leftrightarrow 0\in\partial f(0)+\bigcup_{\lambda\in Y^*_+}\partial(\lambda g)(0)$（参考[107]）.

注 3.1.5 (iii) 给出约束极小问题"min $f(x)$, $g(x)\leqslant 0$"有全局最优解 $x=0$ 的充要条件. 对 3.1.5 (i)、(ii) 可作类似解释.

3.1.6 例（[205]） 设 $u\in X^*$，$f(x)=u(x)+|Bx|$，则 $f:X\to\mathbf{R}$ 是连续凸函数，$\partial f(0)=u+B^*B_1(0)$(2.1.2, 2.1.5, 2.3.2). 由 3.1.5，若 A 满足 3.1.2 中诸条件之一，则 $f(A^{-1}K_-)\geqslant$

$0 \Leftrightarrow u \in A^* K^* + B^* \bar{B}_1(0)$.

由 3.1.2 可得 Krein 定理(1.4.5)的以下推广：

3.1.7 推论 若 $M \subset X$ 是子空间，$K \subset X$ 是闭凸锥，u 是 M 上的连续线性泛函，$u(M \cap K) \geqslant 0$，则 $\exists v \in K^*$，使 $u = v|M$，只要以下条件之一满足：

(i) $\{\lambda|M : \lambda \in K^*\}$ 在 M^* 中弱*闭；

(ii) $\{(x,x) : x \in M\}^{\perp} + 0 \times K^*$ 在 $M^* \times X^*$ 中弱*闭；

(iii) $M + K = X$.

证 取 $A : M \to X$ 为包含映射，则 $A^{-1}K = M \cap K$，$A^*K^* = \{\lambda|M : \lambda \in K^*\}$，于是所要结论由 3.1.2 得出. □

利用 3.1.7，可将 1.4.7(iii) 推广于下.

3.1.8 推论([128]) 设 $K_1, K_2 \subset X$ 是闭凸锥，$K_1 - K_2 = X$，则 $(K_1 \cap K_2)^* = K_1^* + K_2^*$.

证 只需对任给 $u \in (K_1 \cap K_2)^*$ 证 $u \in K_1^* + K_2^*$. 用类似于 1.4.7(ii) 的证法：令 $M = \{(x,x) : x \in X\}$，$K = K_1 \times K_2$，则 M 与 K 分别为 $X \times X$ 中的子空间与闭凸锥. 令 $f(x,x) = u(x)$，则 $f(M \cap K) \geqslant 0$. $\forall x, y \in X$，由 $K_1 - K_2 = X$ 有 $x_1, y_1 \in K_1$，$x_2, y_2 \in K_2$，使得 $x = x_1 - x_2$，$y = y_2 - y_1$，于是

$$(x,y) = -(x_2 + y_1, x_2 + y_1) + (x_1 + y_1, x_2 + y_2) \in M + K,$$

因此 $M + K = X \times X$. 由 3.1.7，有 $\omega = (\lambda, \mu) \in K^* = K_1^* \times K_2^*$，使得 $\omega|M = f$，从而 $u(x) = \lambda(x) + \mu(x) (\forall x \in X)$，即 $u = \lambda + \mu \in K_1^* \times K_2^*$. □

参考文献：[81,93,107,109,112,128,193,205,233].

§3.2 类 凸 性

在转入新的择一定理之前，需要引进一类不同于 §1.5 中所述的广义凸性，它们可通过集的一定凸性来描述.

首先推广凸集概念. 设 $K \subset Y$ 是给定的，以 $\mathscr{N}$ 记空间 Y 的

0-邻域之全体.

3.2.1 定义 若$\forall t\in(0,1),\forall V\in\mathscr{N}:tK+t'K\subset K+V$,则称$K$为**次凸集**.

如所熟知,K凸$\Leftrightarrow\forall t\in(0,1):tK+t'K\subset K$.若将其中的"$\forall t\in(0,1)$"改为"$\exists t\in(0,1)$",则得到稍弱的**近凸集**概念([113]).不过,将3.2.1中的"$\forall t\in(0,1)$"改为"$\exists t\in(0,1)$",得不出更弱的凸性.事实上,"次凸"条件有如下显得更弱的等价形式.

3.2.2 命题 以下条件互相等价:

(i) K为次凸集;

(ii) $\exists t\in(0,1),\forall V\in\mathscr{N}:tK+t'K\subset K+V$;

(iii) $\forall t,\varepsilon\in(0,1),\forall V\in\mathscr{N},\exists r\in(1-\varepsilon,1+\varepsilon):tK+t'K\subset rK+V$;

(iv) $\exists\alpha\in(0,1),\forall\varepsilon\in(0,1),\forall V\in\mathscr{N},\exists r\in(1-\varepsilon,1+\varepsilon):tK+t'K\subset rK+V$;

(v) $\overline{K}$凸.

证 显然(i)$\Rightarrow$(ii)$\Rightarrow$(iv),(i)$\Rightarrow$(iii)$\Rightarrow$(iv).只需证(iv)$\Rightarrow$(v)$\Rightarrow$(i).

证(iv)$\Rightarrow$(v).设α如条件(iv),取定$y_1,y_2\in K$.$\forall\tau=(\varepsilon,V)\in(0,1)\times\mathscr{N}$,取$r_\tau\in(1-\varepsilon,1+\varepsilon),y_\tau\in K,z_\tau\in V$,使$\alpha y_1+\alpha' y_2=r_\tau y_\tau+z_\tau$.因$r_\tau\to1,z_\tau\to0$,故$\alpha y_1+\alpha' y_2=\lim\limits_{\tau}y_\tau\in\overline{K}$.可见$\alpha K+\alpha' K\subset\overline{K}$,因而

$$\alpha\overline{K}+\alpha'\overline{K}\subset\overline{\alpha K+\alpha' K}\subset\overline{K}.$$

令$T=\{t\in J:t\overline{K}+t'\overline{K}\subset\overline{K}\}$,则$T$是闭集,且$\{0,1,\alpha\}\subset T$.若$T\neq J$,则有开区间$(\sigma,\tau)\subset J\backslash T,\sigma,\tau\in T,\sigma<\tau$.令$\beta=\alpha\sigma+\alpha'\tau$,则

$$\begin{aligned}\beta\overline{K}+\beta'\overline{K}&\subset\alpha(\sigma\overline{K}+\sigma'\overline{K})+\alpha'(\tau\overline{K}+\tau'\overline{K})\\&\subset\alpha\overline{K}+\alpha'\overline{K}\subset\overline{K},\end{aligned}$$

可见 $\beta \in T \cap (\sigma, \tau)$，得出矛盾. 因此 $T = J$，这表明 $\overline{K}$ 凸.

证(v) $\Rightarrow$ (i). 设 $\overline{K}$ 凸，取定 $t \in (0,1)$，$V \in \mathscr{N}$，$y_1, y_2 \in K$. 因 $y \stackrel{\text{def}}{=} ty_1 + t'y_2 \in t\overline{K} + t'\overline{K} \subset \overline{K}$，故有 $z \in K \cap (y + V)$. 不妨设 $V = -V$，于是 $y \in K + V$. 可见 $tK + t'K \subset K + V$，即 K 为次凸集. □

显然 K 近凸蕴涵 3.2.2 之条件(ii)，因此有

$$\text{凸} \Rightarrow \text{近凸} \Rightarrow \text{次凸} \Leftrightarrow \text{闭包凸}.$$

以上关系中的"$\Rightarrow$"都是不可逆的.

3.2.3 例 **1°** 设 K 是区间$[0,1]$中有理点之全体，则显然$(K + K)/2 \subset K$，因此 K 近凸而非凸.

2° 设 $K = \{x = (x_1, x_2) \in \mathbf{R}_+^2 : x_1 + x_2 < 1\} \cup \{e_1, e_2\}$，其中 $e_1 = (1,0)$，$e_2 = (0,1)$. 因 $\overline{K} = \{x \in \mathbf{R}_+^2 : x_1 + x_2 \leqslant 1\}$ 凸，故 K 次凸. 另一方面，$\forall t \in (0,1)$，有 $te_1 + t'e_2 \overline{\in} K$，可见 K 不是近凸的.

3.2.4 引理 设 K 近凸，则有 J 的稠子集 T，使得 $\forall t \in T$：$tK + t'K \subset K$.

证 取 $\alpha \in (0,1)$，使得 $\alpha K + \alpha' K \subset K$. 令 $T = \{t \in J : tK + t'K \subset K\}$，则$\{0, 1, \alpha\} \subset T$. $\forall t, s \in T$，令 $\beta = \alpha t + \alpha' s$，则

$$\begin{aligned}\beta K + \beta' K &= (\alpha t + \alpha' s)K + (\alpha t' + \alpha' s')K \\ &\subset \alpha(tK + t'K) + \alpha'(sK + s'K) \\ &\subset \alpha K + \alpha' K \subset K,\end{aligned}$$

可见 $\beta \in T$. 这表明 $\alpha T + \alpha' T \subset T$，即 T 近凸. 于是由 3.2.2 知 $\overline{T}$ 凸，从而 $\overline{T} = J$. □

3.2.5 定理 (i) 若 K 是闭集，则 K 凸 $\Leftrightarrow$ K 近凸 $\Leftrightarrow$ K 次凸.

(ii) 若 K 是开集，则 K 凸 $\Leftrightarrow$ K 近凸.

证 (i) 直接由 3.2.2 推出.

(ii) 设 K 近凸，取 T 如 3.2.4. 若 K 非凸，则 $\exists t_0 \in (0,1)$，x_0，$y_0 \in K$：$z_0 \stackrel{\text{def}}{=} t_0 x_0 + t'_0 y_0 \overline{\in} K$. 令 $\varphi(t) = (z_0 - t' y_0)/t$，则 $\varphi(t_0) = x_0 \in K$. 取 $t_n \in T$：$t_n \to t_0$. 因 K 为开集，φ连续，不妨设 $\varphi(t_n) \in$

K. 于是

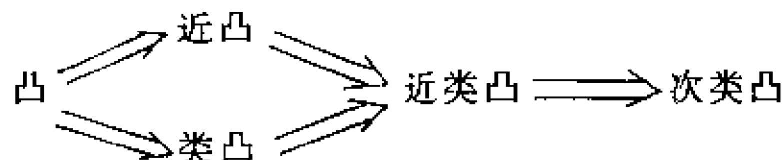
$$z_0 = t_n\varphi(t_n) + t'_n y_0 \in t_n K + t'_n K \subset K,$$

得出矛盾. 可见 K 为凸集. □

现在利用已推广的凸集概念来定义映射 $G: X \to 2^Y$ 的某种广义凸性. 首先回忆起 G 凸 $\Leftrightarrow$ epi G 凸(1.5.1), 因此自然定义 G 近凸 $\Leftrightarrow$ epi G 近凸. 进一步的概念利用集 $R(G) + Y_+$ 的凸性得到.

3.2.6 定义 若 $R(G) + Y_+$ 是凸[近凸或次凸]集, 则说 G 是 Y_+ - **类凸**[Y_+ - **近类凸**或 Y_+ - **次类凸**]的. 若 $-G$ 是 Y_+ - 类凸的, 则说 G 是 Y_+ - **类凹的**; Y_+ - 近类凹等概念仿此. 当不致混淆时, 略去上述记号中的"Y_+ -".

关于映射的上述诸广义凸性的关系如下:



对于 $g: D \to Y$, 可将以上诸广义凸性直接描述如下: g 近凸 $\Leftrightarrow \exists \alpha \in (0,1), \forall x_1, x_2 \in D$, 有

$$g(\alpha x_1 + \alpha' x_2) \leqslant \alpha g(x_1) + \alpha' g(x_2).$$

g 类凸 $\Leftrightarrow \forall t \in (0,1), \forall x_1, x_2 \in D, \exists x \in D$, 使得

$$g(x) \leqslant tg(x_1) + t'g(x_2).$$

g 次类凸 $\Leftrightarrow \forall t \in J, \forall V \in \mathscr{N}, \forall x_1, x_2 \in D, \exists x \in D, y \in V$, 使得

$$g(x) \leqslant tg(x_1) + t'g(x_2) - y.$$

其余仿此. 所宜注意者, 类凸概念不要求定义域 D 具有任何结构!

3.2.7 定理 (i) 若 epi G 闭, 则 G 凸 $\Leftrightarrow$ G 近凸; 若 D 闭且 g 为 $*$ lsc, 则 g 凸 $\Leftrightarrow$ g 近凸; 若 $R(G) + Y_+$ 闭, 则 G 类凸 $\Leftrightarrow$ G 次类凸.

(ii) 设 $Y_+^\circ \neq \varnothing$, 则 G 次类凸 $\Leftrightarrow R(G) + Y_+^\circ$ 凸 $\Leftrightarrow$ G 为 Y_+° 类凸.

证 (i) 关于 G 的结论直接由 3.2.5 (i) 推出. 对于 g 只需指

出：当 D 闭且 g 为 $*$lsc 时，$\mathrm{epi}\, g$ 是闭集. 事实上，若 $x_n \in D, x_n \to x \in D, g(x_n) \leqslant y_n \to y$（这意味着 $(x,y) \in \overline{\mathrm{epi}\, g}$），则 $\forall \lambda \in Y_+^*$，有

$$\lambda g(x) \leqslant \liminf_n \lambda g(x_n) \leqslant \lim_n \lambda(y_n) = \lambda(y).$$

这推出 $g(x) \leqslant y$，从而 $(x,y) \in \mathrm{epi}\, g$.

(ii) 首先设 G 次类凸. 取定 $t \in (0,1), y_i \in R(G), z_i \in Y_+^\circ$ ($i = 1,2$). 因 $z \stackrel{\mathrm{def}}{=} tz_1 + t'z_2 \in Y_+^\circ$，故有 $V \in \mathscr{N}: V + z \subset Y_+^\circ$. 由 $R(G) + Y_+$ 次凸有

$$\begin{aligned}
&t(y_1 + z_1) + t'(y_2 + z_2) \\
&\in tR(G) + t'R(G) + z \\
&\subset t[R(G) + Y_+] + t'[R(G) + Y_+] + z \\
&\subset R(G) + Y_+ + V + z \\
&\subset R(G) + Y_+ + Y_+^\circ \subset R(G) + Y_+^\circ.
\end{aligned}$$

可见 $t[R(G) + Y_+^\circ] + t'[R(G) + Y_+^\circ] \subset R(G) + Y_+^\circ$，即 $R(G) + Y_+^\circ$ 为凸集.

其次设 $R(G) + Y_+^\circ$ 为凸集. 任给 $t \in (0,1), V \in \mathscr{N}$，取 $y_0 \in Y_+^\circ \cap V \cap (-V)$，则

$$\begin{aligned}
&t[R(G) + Y_+] + t'[R(G) + Y_+] \\
&= t[R(G) + Y_+ + y_0] + t'[R(G) + Y_+ + y_0] - y_0 \\
&\subset t[R(G) + Y_+^\circ] + t'[R(G) + Y_+^\circ] + V \\
&\subset R(G) + Y_+^\circ + V \subset R(G) + Y_+ + V.
\end{aligned}$$

可见 $R(G) + Y_+$ 次凸，从而 G 次类凸. □

本节的基本思想源于[191]，但作了较大简化. 亦可参考[93, 159].

§ 3.3 Gordan 定理与 Gale 定理

3.3.1 Gordan 定理 设 $Y_+^\circ \neq \varnothing$，$G$ 为 Y_+-次类凸，则以下

命题两择一：

(P) $\exists y \in R(G): y \ll 0$;

(Q) $\exists \lambda \in Y_+^* \setminus \{0\}: \langle \lambda, R(G) \rangle \geqslant 0 \ (\Leftrightarrow \lambda \in R(G)^*)$.

若 $D(G) = X, D(G^*) = Y^*$，则(Q)等价于命题

(Q′) $\exists \lambda \in Y_+^* \setminus \{0\}: G^* \lambda = 0$.

证 若(P)无解，则 $0 \overline{\in} R(G) + Y_+^\circ \overset{\text{def}}{=} C$. C 必为开凸集(3.2.7)，故由1.1.2有 $\lambda \in Y^* \setminus \{0\}: 0 \leqslant \langle \lambda, C \rangle$. 这推出 $\lambda \in C^* = R(G)^* \cap Y_+^*$，即(Q)成立.

设 $D(G) = X, D(G^*) = Y^*, \lambda \in Y^*$. 显然从 $G^* \lambda = 0$ 推出 $\lambda \in R(G)^*$. 反之，若 $\lambda \in R(G)^*$，则对任给 $u \in G^*(-\lambda)$，有 $-u(x) - \lambda(Gx) \geqslant 0 \ (\forall x \in X)$，从而 $-u(x) \geqslant \lambda(Gx) \geqslant 0$，这推出 $u = 0$. 因 $D(G^*) = Y^*$，故必 $G^*(-\lambda) = 0$. 这又推出 $-\lambda \in R(G)^*$，因此 $G^* \lambda = 0$. □

3.3.2 推论 设 $Y_+^\circ \neq \varnothing$，$g$ 为 Y_+-次类凸映射，则以下命题两择一：

(P) $\exists x \in D: g(x) \ll 0$;

(Q) $\exists \lambda \in Y_+^* \setminus \{0\}: \langle \lambda, g(D) \rangle \geqslant 0 (\Leftrightarrow \lambda \in g(D)^*)$.

关于3.3.2可参考[85,104,105].

看一简单例子. 设 $D = \{x \in \mathbf{R}_+^2 : x_1 + x_2 > 1\} \cup \{(1,0), (0,1)\}$，则 $g = \mathrm{id} : D \to \mathbf{R}^2$ 为 $\mathbf{R}_+^2$-次类凸(但非类凸)，问题"$\exists x \in D : g(x) = x \ll 0$"无解. 由3.3.2，问题

$$\exists \lambda \in \mathbf{R}_+^2 \setminus \{0\}: \langle \lambda, g(D) \rangle = \langle \lambda, D \rangle \geqslant 0$$

应有解. 事实上，$\lambda = (1,1)$ 就是这样的解.

3.3.3 推论(线性Gordan定理) 设 $Y_+^\circ \neq \varnothing$，$G$ 是线性的，$D(G) = X, D(G^*) = Y^*$，则3.3.1中的命题(P)与(Q′)两择一. 特别，以下命题两择一：

(P) $\exists x \in X: Ax \ll 0$;

(Q) $\exists \lambda \in Y_+^* \setminus \{0\}: A^* \lambda = 0$.

将3.3.3用到$A \in \mathbf{R}^{m\times n} \cong L(\mathbf{R}^n,\mathbf{R}^m)$，得出以下命题两择一：

(P) $\exists x \in \mathbf{R}^n : Ax \ll 0$;

(Q) $\exists y \in \mathbf{R}^m_+ \setminus \{0\} : A^T y = 0$.

经典的 Gordan 定理乃指此而言. 3.3.1 大概是这一结果迄今所知的最一般的推广.

对于 3.3.1 ~ 3.3.3，条件 $Y^*_+ \neq \varnothing$ 显然是不可缺少的. 本节的后半部分无需这一假设.

3.3.4 Gale 定理 假定：(i) G 为 Y_+ - 次类凸；(ii) 存在 Y 的 0-邻域V，使$C \cap V$ 为闭集，$C = R(G) + Y_+$. 则以下命题两择一：

(P) $\exists y \in R(G) : y \leqslant 0$;

(Q) $\exists \lambda \in Y^*_+ : \langle \lambda, R(G) \rangle > 0$.

条件(ii) 可代以如下条件之一：(iii) epi G 闭；存在 Y 的闭 0-邻域 V，使 $G^{-1}(V - Y_+)$ 紧. (iv) X 为自反 B-空间，当 X 中用弱拓扑时 epi G 为闭集；$\exists \bar{\lambda} \in Y^*_+$，使

$$\liminf_{y \in Gx, |x| \to \infty} \bar{\lambda}(y) > 0. \tag{1}$$

当 D_G 有界时认定(1) 自动成立(下文中在类似情况下概如此).

证 设(P) 无解，则 $0 \overline{\in} \bar{C}$. 否则，有 C 中的网$\{c_\alpha\} : c_\alpha \to 0$. 不妨设$\{c_\alpha\} \subset V$，于是由条件(ii) 有 $0 \in C$，这与(P) 无解矛盾. 由条件(i) 与 3.2.6，3.2.2，$\bar{C}$ 为闭凸集，于是由 1.1.2 有 $\lambda \in Y^*$，使 $0 < \langle \lambda, C \rangle = \langle \lambda, R(G) \rangle + \langle \lambda, Y_+ \rangle$，这推出 $\lambda \in Y^*_+$ 且$\langle \lambda, R(G) \rangle > 0$.

证(iii) $\Rightarrow$ (ii). 设 V 如条件(iii)，今证 $C \cap V$ 为闭集. 设$\{z_\alpha\}$ 是集 $C \cap V$ 中的网，$z_\alpha \to z$，则 $z \in V$. 由 $z_\alpha \in C$ 有 $x_\alpha \in X$，$y_\alpha \in Gx_\alpha$，使得 $y_\alpha \leqslant z_\alpha$. 因 $y_\alpha = z_\alpha - (z_\alpha - y_\alpha) \in V - Y_+$，故 $x_\alpha \in G^{-1}(V - Y_+)$. 由 $G^{-1}(V - Y_+)$ 紧，不妨设 $x_\alpha \to x$. 因$(x_\alpha, z_\alpha) \in$ epi G 且 epi G 闭，故$(x, z) \in$ epi G，从而 $z \in C \cap V$，$C \cap V$ 为闭集得证.

证(iv) $\Rightarrow$ (ii). 由条件(iv) 有 $\varepsilon > 0$，使当 $|x| \to \infty$ 时，$\forall y \in Gx$，有 $\bar{\lambda}(y) > \varepsilon$. 取 Y 的闭 0-邻域 V，使 $\bar{\lambda}(v) \subset [-\varepsilon, \varepsilon]$. 设 $z \in$

$C \cap V$. 取 $z_n \in C \cap V$，使 $z_n \to z$；取 $x_n \in X, y_n \in Gx_n$，使 $y_n \leqslant z_n$. 则 $\bar{\lambda}(y_n) \leqslant \bar{\lambda}(z_n) \leqslant \varepsilon$，故 $\{x_n\}$ 有界. 由 1.1.6，不妨设 $x_n \rightharpoonup x$. 因 $(x_n, z_n) \in \text{epi}\, G$ 且 epi G 闭，故 $(x,z) \in \text{epi}\, G$，因此 $z \in C \cap V$，$C \cap V$ 为闭集. □

以 g 代 G 从 3.3.4 得出：

3.3.5 推论 假定：(i) g 为 Y_+ - 次类凸；(ii) 存在 Y 的 0- 邻域 V，使 $(g(D) + Y_+) \cap V$ 为闭集. 则以下命题两择一：

(P) $\exists x \in D: g(x) \leqslant 0$；

(Q) $\exists \lambda \in Y_+^*: \langle \lambda, g(D) \rangle > 0$.

条件(ii) 可代以如下条件之一：

(iii) epi g 闭；存在 Y 的闭 0 邻域 V，使 $g^{-1}(V - Y_+)$ 紧.

(iv) X 为自反 B- 空间，D 弱闭且 g 为 $*$ wlsc；

$$\exists \lambda \in Y_+^*: \liminf_{x \in D, \|x\| \to \infty} \lambda g(x) > 0. \tag{2}$$

证 只需指出：D 弱闭且 g 为 $*$ wlsc 推出，当 X 中用弱拓扑时 epi g 为闭集. 设 $(x_n, y_n) \in \text{epi}\, g$，$x_n \rightharpoonup x$，$y_n \to y$，则 $x \in D$，且 $\forall \lambda \in Y_+^*$，于是有

$$\lambda g(x) \leqslant \liminf_{n \to \infty} \lambda g(x_n) \leqslant \lim_n \lambda(y_n) = \lambda(y),$$

这推出 $g(x) \leqslant y$. 因此 $(x,y) \in \text{epi}\, g$，从而 epi g 为闭集. □

将 3.3.5 用到 $g = A - a(a \in Y)$ 上，得

3.3.6 推论(线性 Gale 定理) 若 $R(A) + Y_+$ 为闭集，则以下命题两择一：

(P) $\exists x \in X: Ax \leqslant a$；

(Q) $\exists \lambda \in Y_+^*: A^* \lambda = 0, \lambda(a) = -1$.

将 3.3.6 用到 $A \in \mathbf{R}^{m \times n}, a \in \mathbf{R}^m$，得出以下命题两择一：

(P) $\exists x \in \mathbf{R}^n: Ax \leqslant a$；

(Q) $\exists y \in \mathbf{R}_+^m: A^{\mathrm{T}} y = 0, \langle y, a \rangle = -1$.

经典的 Gale 定理原本指此.

若以 $Y_+ \times 0 (\subset Y \times Z)$ 取代 Y_+，则从 3.3.4 得到 Gale 定理

的如下推广形式：

3.3.7 推论 假定：(i) $\Phi \stackrel{\text{def}}{=} (G,H)$ 为 $Y_+ \times 0$-次类凸；(ii) 存在 $Y \times Z$ 的 0-邻域 V，使 $C \cap V$ 为闭集，$C = R(\Phi) + Y_+ \times 0$，则以下命题两择一：

(P) $\exists (y,z) \in R(\Phi): y \leqslant 0, z = 0$;

(Q) $\exists (\lambda,\mu) \in Y_+^* \times Z^*: R(\lambda G + \mu H) > 0$.

3.3.5 与 3.3.6 亦有相应的推广：

3.3.8 推论 假定：(i) $\varphi \stackrel{\text{def}}{=} (g,h)$ 为 $Y_+ \times 0$-次类凸；(ii) 存在 $Y \times Z$ 的 0-邻域 V，使 $C \cap V$ 为闭集，$C = \varphi(D) + Y_+ \times 0$. 则以下命题两择一：

(P) $\exists x \in D: g(x) \leqslant 0, h(x) = 0$;

(Q) $\exists (\lambda,\mu) \in Y_+^* \times Z^*: (\lambda g + \mu h)(D) > 0$.

条件(ii) 可代以如下条件之一：

(iii) $\operatorname{epi} \varphi = \{(x,y,h(x)): g(x) \leqslant y\}$ 闭；存在 Y,Z 的闭 0-邻域 U,V，使 $g^{-1}(U - Y_+) \cap h^{-1}(V)$ 紧.

(iv) X 为自反 B-空间，D 弱闭，g 为 $*$ wlsc，在 D 中 $x_n \rightharpoonup x \Rightarrow h(x_n) \rightharpoonup h(x)$；且 $\exists (\lambda,\mu) \in Y_+^* \times Z^*$，使得

$$\liminf_{x \in D, |x| \to \infty} \lambda g(x) + \mu h(x) > 0. \tag{3}$$

3.3.9 推论 设 $a \in Y, b \in Z, R(A,B) + Y_+ \times 0$ 为闭集，则以下命题两择一：

(P) $\exists x \in X: Ax \leqslant a, Bx = b$;

(Q) $\exists (\lambda,\mu) \in Y_+^* \times Z^*: A^*\lambda + B^*\mu = 0, \lambda(a) + \mu(b) = -1$.

参考文献：[73,85,93,104,105].

§ 3.4 Motzkin 定理

本节设 $W_+ \neq \varnothing \neq Y_+$，以 $\mathscr{N}_X$ 记 X 的 0-邻域之集，$\mathscr{N}_Y$ 等

仿此.

3.4.1 定理([95]) 设$\Phi=(F,G,H)$满足以下条件:(i) $\forall t\in J,\forall U\in\mathscr{N}_W,\forall V\in\mathscr{N}_Y$: $tR(\Phi)+t'R(\Phi)\subset R(\Phi)+W_+\times Y_+\times 0+U\times V\times 0$;(ii) $\exists(x_0,w_0,y_0,z_0)\in\mathrm{Gr}\,\Phi,\forall U\in\mathscr{N}_W,\forall V\in\mathscr{N}_Y,\exists N\in\mathscr{N}_X,\forall x\in D_F\cap D_G\cap(x_0+N)$: $Fx\cap(w_0+U-W_+)\neq\varnothing\neq Gx\cap(y_0+V-Y_+)$;(iii) $\forall N\in\mathscr{N}_X,\exists B\in\mathscr{N}_Z$: $z_0+B\subset H(D_F\cap D_G\cap(x_0+N))$;(iv) $\exists(\hat{x},\hat{w},\hat{y},0)\in\mathrm{Gr}\,\Phi$: $\hat{y}\ll 0,R(H)^*=\{0\}$. 则以下命题两择一:

(P) $\exists(w,y,z)\in R(\Phi)$: $w\ll 0,y\leqslant 0,z=0$;

(Q) $\exists(\rho,\lambda,\mu)\in(W_+^*\backslash\{0\})\times Y_+^*\times Z^*$:

$$R(\rho F+\lambda G+\mu H)\geqslant 0.$$

当W,Y,Z为有限维空间时不必用到条件(ii)(iii).

证 设(P)无解,则$0\bar{\in}R(\Phi)+W_+^\circ\times Y_+^\circ\times 0\overset{\text{def}}{=}C$. 令$K=W_+^\circ\times Y_+^\circ\times 0$. $\forall k\in K$,取$U\in\mathscr{N}_W$,$V\in\mathscr{N}_Y$,使$k+U\times V\times 0\subset K$. 于是由条件(i),$\forall t\in J$,有

$$\begin{aligned}tR(\Phi)+t'R(\Phi)+k&\subset R(\Phi)+K+U\times V\times 0+k\\&\subset R(\Phi)+K+K\subset R(\Phi)+K=C.\end{aligned}$$

这推出$tC+t'C\subset C(\forall t\in J)$,即$C$为凸集. 取$\bar{w}\gg w_0,\bar{y}\gg y_0$,设$(w,y,z)\approx(\bar{w},\bar{y},z_0)$. 取$U\in\mathscr{N}_W$,$V\in\mathscr{N}_Y$,使$w_0+U\ll w,y_0+V\ll y$. 设$N,B$分别如条件(ii),(iii). 不妨设$z\in z_0+B$,于是有$x\in D_F\cap D_G\cap(x_0+N)$: $z\in Hx$. 取$w'\in Fx\cap(w_0+U-W_+)$,$y'\in Gx\cap(y_0+V-Y_+)$,则

$$\begin{aligned}w-w'&\in w-w_0-U+W_+\\&\subset W_+^\circ+W_+\subset W_+^\circ.\end{aligned}$$

同理$y-y'\in Y_+^\circ$. 于是

$$\begin{aligned}(w,y,z)&\in\Phi x+(w-w',y-y',0)\\&\subset R(\Phi)+W_+^\circ\times Y_+^\circ\times 0=C.\end{aligned}$$

这表明 $(\overline{w},\overline{y},z_0)\in C\neq\varnothing$. 由 1.1.2，有 $0\neq\varphi=(\rho,\lambda,\mu)\in W^*\times Y^*\times Z^*$，使 $0\leqslant\langle\varphi,C\rangle=\langle\varphi,R(\Phi)\rangle+\langle\varphi,K\rangle$. 这推出

$$\varphi\in K^*=W_+^*\times Y_+^*\times Z^*,$$

$$0\leqslant\langle\varphi,R(\Phi)\rangle=R(\rho F+\lambda G+\mu H).$$

必 $\rho\neq 0$. 否则，由条件(iv) 有

$$0\leqslant\langle\varphi,(\hat{w},\hat{y},0)\rangle=\lambda(\hat{y})\leqslant 0,$$

这推出 $\lambda=0, 0\neq\mu\in R(H)^*$，得出矛盾.

若 W,Y,Z 为有限维空间，则可用 1.1.2(iii) 而不必验证 $C\neq\varnothing$，因此不必用条件(ii)，(iii). □

注 3.4.1 的条件(i) 弱于"Φ 为 $W_+\times Y_+\times 0$-类凸"，但强于"Φ 为 $W_+\times Y_+\times 0$-次类凸".

3.4.2 推论(非线性 Motzkin 定理) 假定：(i) $\forall t\in J,\forall U\in\mathscr{N}_W,\forall V\in\mathscr{N}_Y$ ：$t\varphi(D)+t'\varphi(D)\subset\varphi(D)+W_+\times Y_+\times 0+U\times V\times 0,\varphi=(f,g,h)$；(ii)$\exists x_0\in D,\forall U\in\mathscr{N}_W,\forall V\in\mathscr{N}_Y,\exists N\in\mathscr{N}_X$ ：$f(D\cap(x_0+N))\subset f(x_0)+U-W_+,g(D\cap(x_0+N))\subset g(x_0)+V-Y_+$；(iii) $\forall N\in\mathscr{N}_X$，$\exists B\in\mathscr{N}_Z$ ：$h(x_0)+B\subset h(D\cap(x_0+N))$；(iv) $\exists\hat{x}\in D$：$g(\hat{x})\ll 0,h(\hat{x})=0$；$h(D)^*=\{0\}$. 则以下命题两择一：

(P) $\exists x\in D$：$f(x)\ll 0,g(x)\leqslant 0,h(x)=0$；

(Q) $\exists(\rho,\lambda,\mu)\in(W_+^*\setminus\{0\})\times Y_+^*\times Z^*$：

$$(\rho f+\lambda g+\mu h)(D)\geqslant 0.$$

若 W,Y,Z 为有限维空间，则不必用条件(ii)，(iii).

注 当 3.4.2 中条件(ii) 满足时，约定说"f 与 g 在 x_0 处为上半连续(usc)". 对于数量函数，这就是通常的 usc. 显然，f 连续 $\Rightarrow$ f 为 usc $\Rightarrow$ f 为 $*$usc，后者意味着 $\forall\rho\in W_+^*$：ρf 为 usc.

若以 T,A,B 取代 f,g,h，则 3.4.2 的条件(i)，(ii) 自动满足，当 $R(B)=Z$ 时条件(iii) 亦满足且 $R(B)^*=R(B)^\perp=\{0\}$. 于是有：

3.4.3 推论(线性 Motzkin 定理)　假定:(i) $R(B) = Z$;(ii) $\exists \hat{x} \in N(B): A\hat{x} \ll 0$. 则以下命题两择一:

(P) $\exists x \in N(B): Tx \ll 0, Ax \leqslant 0$;

(Q) $\exists (\rho, \lambda, \mu) \in (W_+^* \backslash \{0\}) \times Y_+^* \times Z^*$:

$$T^* \rho + A^* \lambda + B^* \mu = 0.$$

注　若 W, Y, Z 为有限维空间,则在 3.4.3 中无需假定(i),(ii),可从(P)无解推出:$\exists (\rho, \lambda, \mu)x \in (W_+^* \times Y_+^* \times Z^*) \backslash \{0\}$: $T^* \rho + A^* \lambda + B^* \mu = 0$.

取 $Y = \{0\}$ 从 3.4.1 得出:

3.4.4 推论　假定:(i) $\forall t \in J, \forall U \in \mathscr{N}_W: tR(\Phi) + t'R(\Phi) \subset R(\Phi) + W_+ \times 0 + U \times 0, \Phi = (F, H)$;(ii) $\exists (x_0, w_0, z_0) \in \mathrm{Gr}\, \Phi, \forall U \in \mathscr{N}_W, \exists N \in \mathscr{N}_X, \forall x \in D_F \cap (x_0 + N): Fx \cap (w_0 + U - W_+) \neq \varnothing$;(iii) $\forall N \in \mathscr{N}_X, \exists B \in \mathscr{N}_Z: z_0 + B \subset H(D_F \cap (x_0 + N))$;(iv) $R(H)^* = \{0\}$. 则以下命题两择一:

(P) $\exists (w, z) \in R(\Phi): w \ll 0, z = 0$;

(Q) $\exists (\rho, \mu) \in (W_+^* \backslash \{0\}) \times Z^*: R(\rho F + \mu H) \geqslant 0$.

若 W, Z 为有限维空间,则不必用条件(ii),(iii).

3.4.5 推论　假定:(i) $\forall t \in J, \forall U \in \mathscr{N}_W: t\varphi(D) + t'\varphi(D) \subset \varphi(D) + W_+ \times 0 + U \times 0, \varphi = (f, h)$;(ii) $\exists x_0 \in D: f$ 在 x_0 为 usc;(iii) $\forall N \in \mathscr{N}_X, \exists B \in \mathscr{N}_Z: h(x_0) + B \subset h(D \cap (x_0 + N))$;(iv) $h(D)^* = \{0\}$. 则以下命题两择一:

(P) $\exists x \in D: f(x) \ll 0$;

(Q) $\exists (\rho, \mu) \in (W_+^* \backslash \{0\}) \times Z^*: (\rho f + \mu h)(D) \geqslant 0$.

特别,若 $R(B) = Z, a \in W, b \in Z$,则命题"$\exists x \in X: Tx \ll a, Bx = b$"与"$\exists (\rho, \mu) \in (W_+^* / \{0\}) \times Z^*: T^* \rho + B^* \mu = 0, \rho(a) + \mu(b) \leqslant 0$"两择一.

以 f 代 F,从 3.4.4 得:

3.4.6 推论 假定：(i) $\forall t \in J, \forall U \in \mathcal{N}_W: tR(\Phi) + t'R(\Phi) \subset R(\Phi) + W_+ \times 0 + U \times 0, \Phi = (f, H)$；(ii) $\exists x_0 \in D$：f 在 x_0 为 usc；(iii) $\exists z_0 \in Hx_0, \forall N \in \mathcal{N}_X, \exists B \in \mathcal{N}_Z: z_0 + B \subset H(D \cap (x_0 + N))$；(iv) $R(H)^* = \{0\}$. 则以下命题两择一：

(P) $\exists x \in D: f(x) \ll 0, 0 \in Hx$；

(Q) $\exists (\rho, \mu) \in (W_+^* \backslash \{0\}) \times Z^*: R(\rho f + \mu H) \geqslant 0$.

若 $Z = \{0\}$，则相应的结果较为简单，且可直接从 3.3.1 推出而不必用到 3.4.1.

3.4.7 定理 假定：(i) $\Phi \overset{\text{def}}{=} (F, G)$ 为 $W_+ \times Y_+$ - 次类凸；(ii) $\exists (\hat{w}, \hat{y}) \in R(\Phi): \hat{y} \ll 0$. 则以下命题两择一：

(P) $\exists (w, y) \in R(\Phi): w \ll 0, y \leqslant 0$；

(Q) $\exists (\rho, \lambda) \in (W_+^* \backslash \{0\}) \times Y_+^*: R(\rho F + \lambda G) \geqslant 0$.

证 若(P)无解，则问题"$\exists u \in R(\Phi): u \ll 0$"无解. 于是由 3.3.1 有 $0 \neq (\rho, \lambda) \in W_+^* \times Y_+^*: R(\rho F + \lambda G) \geqslant 0$. 由条件(ii)推出 $\rho \neq 0$. □

3.4.8 推论 假定：(i) (f, g) 为 $W_+ \times Y_+$ - 次类凸；(ii) $\exists \hat{x} \in D: g(\hat{x}) \ll 0$，则以下命题两择一：

(P) $\exists x \in D: f(x) \ll 0, g(x) \leqslant 0$；

(Q) $\exists (\rho, \lambda) \in (W_+^* \backslash \{0\}) \times Y_+^*: (\rho f + \lambda g)(D) \geqslant 0$.

特别，若 $\exists \hat{x} \in X: A\hat{x} \ll 0$，则命题"$\exists x \in X: Tx \ll 0, Ax \leqslant 0$"与"$\exists (\rho, \lambda) \in (W_+^* \backslash \{0\}) \times Y_+^*: T^*\rho + A^*\lambda = 0$"两择一.

3.4.8 推广了[72, Th. 3.1]与[104, Th. 5.1].

对于 3.4.1 及其推论，条件"$Y_+^* \neq \varnothing$"是重要的. 倘无此条件，某种变形的 Motzkin 定理仍可建立.

3.4.9 定理 假定：(i) $\Phi \overset{\text{def}}{=} (F, G)$ 为 $W_+ \times Y_+$ - 次类凸；(ii) 存在 $w_0 \in W_+$ 与 Y 的 0- 邻域 V，使 $C \cap ((w_0 - W_+) \times V)$ 为闭集，$C = R(\Phi) + W_+ \times Y_+$，则以下命题两择一：

(P) $\exists (w, y) \in R(\Phi): w \ll 0, y \leqslant 0$；

(Q) $\forall w \in W_-^\circ, \exists (\rho,\lambda) \in W_+^* \times Y_+^*$:

$$R(\rho F + \lambda G) > \rho(w).$$

若假定：(iii) $\exists (\hat{w},\hat{y}) \in R(\Phi)$; $\hat{y} \leqslant 0$，则可要求(Q) 中的 $\rho \neq 0$；当 $W = \mathbf{R}$ 时可取 $\rho = 1$. 条件(ii) 可代以如下条件之一：(iv) epi Φ 闭；存在 $w_0 \in W_+$ 与 Y 的闭 0- 邻域 V，使 $F^{-1}(w_0 - W_+) \cap G^{-1}(V - Y_+)$ 紧. (v) X 为自反 B- 空间，当 X 中用弱拓扑时 epi Φ 闭，且 $\exists (\bar{\rho},\bar{\lambda}) \in W_+^* \times Y_+^*$，使

$$\liminf_{(w,y) \in \Phi x, |x| \to \infty} \bar{\rho}(w) + \bar{\lambda}(y) > 0.$$

证 设 (P) 无解，取定 $w \in W_-^\circ$. 必定 $(w,0) \overline{\in} \bar{C}$. 否则，有网 $\{c_\alpha\} \subset C$; $c_\alpha \to (w,0)$. 因 $w \in W^\circ \subset (w_0 - W_+)^\circ$，故可设 $c_\alpha \in (w_0 - W_+) \times V$，于是由条件(ii) 有 $(w,0) \in C$，这矛盾于(P) 无解. 由条件(i) 及 3.2.6, 3.2.2 知 C 为闭凸集. 于是由 1.1.2 有 $\tau = (\rho,\lambda) \in W^* \times Y^*$，使得

$$\rho(w) = \langle \tau, (w,0) \rangle < \langle \tau, C \rangle$$
$$= \langle \tau, R(\Phi) \rangle + \langle \tau, W_+ \times Y_+ \rangle.$$

由此推出 $\tau \in W_+^* \times Y_+^*$, $\rho(w) < \langle \tau, R(\Phi) \rangle = R(\rho F + \lambda G)$. 这表明(Q) 有解.

若条件(iii) 满足，$(0,\lambda) \in W_+^* \times Y_+^*$ 满足(Q)，则 $0 < \lambda(\hat{y}) \leqslant 0$，得出矛盾. 故可设(Q) 中 $\rho \neq 0$.

证(iv) $\Rightarrow$ (ii). 设 w_0, V 如条件(iv)，今证 $C \cap ((w_0 - W_+) \times V)$ 闭. 设 $\{(w_\alpha, y_\alpha)\}$ 是 $C \cap ((w_0 - W_+) \times V)$ 中的网，$(w_\alpha, y_\alpha) \to (w,y)$，则 $(w,y) \in (w_0 - W_+) \times V$. 取 $x_\alpha \in X$, $(\overline{w}_\alpha, \overline{y}_\alpha) \in \Phi x_\alpha$，使 $\overline{w}_\alpha \leqslant w_\alpha$, $\overline{y}_\alpha \leqslant y_\alpha$，则 $\overline{w}_\alpha \in w_0 - W_+$, $\overline{y}_\alpha \in V - Y_+$，于是 $x_\alpha \in F^{-1}(w_0 - W_+) \cap G^{-1}(V - Y_+)$. 由条件(iv)，不妨设 $x_\alpha \to x$. 由 $(x_\alpha, w_\alpha, y_\alpha) \in$ epi Φ 及 epi Φ 闭，有 $(x,w,y) \in$ epi Φ，从而 $(w,y) \in C$. 因此 $C \cap ((w_0 - W_+) \times V)$ 为闭集.

证(v)$\Rightarrow$(ii). 由条件(v) 有 $\varepsilon > 0$，使当 $|x| \to \infty$ 时 $\forall w \in Fx$, $y \in Gx$，有 $\bar{\rho}(w) + \bar{\lambda}(y) > 2\varepsilon$. 取 $w_0 \in W_+^\circ$，使 $\bar{\rho}(w_0) < \varepsilon$；

取 Y 的闭 0- 邻域 V，使 $\bar{\lambda}(V) \subset [-\epsilon,\epsilon]$. 今证 $C \cap ((w_0 - W_+) \times V)$ 闭. 设 $(w_n, y_n) \in C \cap ((w_0 - W_+) \times V)$，$(w_n, y_n) \to (w, y)$. 取 $x_n \in X$，$\overline{w}_n \in Fx_n$，$\bar{y}_n \in Gx_n$，使 $\overline{w} \leqslant w_n$，$\bar{y}_n \leqslant y_n$. 因

$$\bar{\rho}(\overline{w}_n) + \bar{\lambda}(\bar{y}_n) \leqslant \bar{\rho}(w_0) + \bar{\lambda}(y_n) \leqslant 2\epsilon,$$

故 $\{x_n\}$ 有界，因此不妨设 $x_n \rightharpoonup x$. 于是 $(x, w, y) \in \mathrm{epi}\ \Phi$，从而 $(w, y) \in C \cap (W_- \times V)$，$C \cap (W_- \times V)$ 为闭集. □

注 在证(v) ⇒ (ii) 时指明了 $w_0 \in W_+$，这一点在后面（如 5.2.8）将用到.

以 (f, g) 代 (F, G) 从 3.4.9 得到：

3.4.10 推论 假定：(i) $\varphi \overset{\mathrm{def}}{=} (f, g)$ 为 $W_+ \times Y_+$ - 次类凸；(ii) 存在 $w_0 \in W_+$ 与 Y 的 0- 邻域 V，使 $C \cap ((w_0 - W_+) \times V)$ 闭，$C = \varphi(D) + W_+ \times Y_+$. 则以下命题两择一：

(P) $\exists x \in D: f(x) \ll 0, g(x) \leqslant 0$；

(Q) $\forall w \in W_-^\circ$，$\exists (\rho, \lambda) \in W_+^* \times Y_+^*$：

$$(\rho f + \lambda g)(D) > \rho(w).$$

若假定：(iii) $\exists \hat{x} \in D: g(\hat{x}) \leqslant 0$，则可要求(Q) 中 $\rho \neq 0$. 条件(ii) 可代以如下条件之一：(iv) epi φ 闭；存在 $w_0 \in W_+$ 与 Y 的闭 0- 邻域 V，使 $f^{-1}(w_0 - W_+) \cap g^{-1}(V - Y_+)$ 紧. (v) X 为自反 B 空间，D 弱闭，f 与 g 为 $*$ wlsc；存在 $(\rho, \lambda) \in W_+^* \times Y_+^*$，使

$$\liminf_{x \in D, \|x\| \to \infty} \rho f(x) + \lambda g(x) > 0.$$

证 仅需指出：在条件(v) 下，当 X 中用弱拓扑时 epi φ 闭. 设 $(x_n, w_n, y_n) \in \mathrm{epi}\ \varphi$，$x_n \rightharpoonup x$，$w_n \to w$，$y_n \to y$. 则 $x \in D$，且 $\forall \rho \in W_+^*$ 有

$$\rho f(x) \leqslant \liminf_{n \to \infty} \rho f(x_n) \leqslant \lim_n \rho(w_n) = \rho(w).$$

这推出 $f(x) \leqslant w$；同理 $g(x) \leqslant y$，因此 $(x, w, y) \in \mathrm{epi}\ \varphi$. □

以 $Y_+ \times 0 \subset Y \times Z$ 代 Y_+，从 3.4.9，3.4.10 得出：

3.4.11 推论 假定：(i) $\Phi \overset{\mathrm{def}}{=} (F, G, H)$ 为 $W_+ \times Y_+ \times 0$- 次类凸；(ii) 存在 $w_0 \in W_+$ 与 $Y \times Z$ 的 0- 邻域 V，使 $C \cap ((w_0 -$

$W_+) \times V)$ 闭，$C = R(\Phi) + W_+ \times Y_+ \times 0$. 则以下命题两择一：

(P) $\exists (w,y,z) \in R(\Phi): w \ll 0, y \leqslant 0, z = 0$;

(Q) $\forall w \in W_-^\circ, \exists (\rho,\lambda,\mu) \in W_+^* \times Y_+^* \times Z^*$:

$$R(\rho F + \lambda G + \mu H) > \rho(w).$$

若假定：(iii) $\exists (\hat{w},\hat{y},\hat{z}) \in R(\Phi): \hat{y} \leqslant 0, \hat{z} = 0$，则可要求(Q)中的 $\rho \neq 0$. 条件(ii)可代以如下条件之一：(iv) epi $\Phi \stackrel{\text{def}}{=} \{(x, w,y,z): \exists (\overline{w},\overline{y},z) \in \Phi x$，使 $\overline{w} \leqslant w, \overline{y} \leqslant y\}$ 闭；存在 $w_0 \in W_+$ 与 Y,Z 的闭 0-邻域 U,V，使 $F^{-1}(w_0 - W_+) \cap G^{-1}(U - Y_+) \cap H^{-1}(V)$ 紧. (v) X 为自反 B-空间，当 X 中用弱拓扑时 epi Φ 闭，且存在 $(\rho,\lambda,\mu) \in W_+^* \times Y_+^* \times Z^*$，使

$$\liminf_{(w,y,z) \in \Phi x, |x| \to \infty} \rho(w) + \lambda(y) + \mu(z) > 0.$$

3.4.12 推论 假定：(i) $\varphi \stackrel{\text{def}}{=} (f,g,h)$ 为 $W_+ \times Y_+ \times 0$-次类凸；(ii) 存在 $w_0 \in W_+$ 与 $Y \times Z$ 的 0-邻域 V，使 $C \cap ((w_0 - W_+) \times V)$ 闭，$C = \varphi(D) + W_+ \times Y_+ \times 0$. 则以下命题两择一：

(P) $\exists x \in D: f(x) \ll 0, g(x) \leqslant 0, h(x) = 0$;

(Q) $\forall w \in W_-^\circ, \exists (\rho,\lambda,\mu) \in W_+^* \times Y_+ \times Z^*$:

$$(\rho f + \lambda g + \mu h)(D) > \rho(w).$$

若假定：(iii) $\exists \hat{x} \in D: g(\hat{x}) \leqslant 0, h(\hat{x}) = 0$，则可设(Q)中的 $\rho \neq 0$；若进而设 $W = \mathbf{R}$，则(Q)可改写成：

$$\forall \varepsilon > 0, \exists (\lambda,\mu) \in Y_+^* \times Z^*: (f + \lambda g + \mu h)(D) > -\varepsilon.$$

条件(ii)可代以如下条件之一：(iv) epi $\varphi \stackrel{\text{def}}{=} \{(x,w,y, h(x)): f(x) \leqslant w, g(x) \leqslant y\}$ 闭；存在 $w_0 \in W_+$ 与 Y,Z 的闭 0-邻域 U,V，使 $f^{-1}(w_0 - W_+) \cap g^{-1}(U - Y_+) \cap h^{-1}(V)$ 紧. (v) X 为自反 B-空间，D 弱闭，f 与 g 为 $*$wlsc，在 D 中 $x_n \rightharpoonup x \Rightarrow h(x_n) \rightharpoonup h(x)$，存在 $(\rho,\lambda,\mu) \in W_+^* \times Y_+^* \times Z^*$，使

$$\liminf_{x \in D, |x| \to \infty} \rho f(x) + \lambda g(x) + \mu h(x) > 0.$$

以下这个较特殊的结果在§6.2中将要用到.

3.4.13 定理(Tucker) 设 X,Y,W 是有限维空间，则以下命

题两择一：

(P) $\exists x \in X: Tx < 0, Ax \leqslant 0$；

(Q) $\exists (\rho, \lambda) \in \mathring{W}_+^* \times Y_+^*: T^*\rho + A^*\lambda = 0$.

证 令 $S = (T, A)$. 若(Q)无解，则 $(\mathring{W}_+^* \times Y_+^* \cap N(S^*) = \varnothing$. 取含 $N(S^*)$ 的超平面 $\pi = N(\varphi), \varphi = (w, y) \in (W \times Y)^{**} = W \times Y$，使得 $\pi \cap (\mathring{W}_+^* \times Y_+^*) = \varnothing$. 可设 $\langle \mathring{W}_+^* \times Y_+^*, \varphi \rangle < 0$，因此 $-\varphi \in (\mathring{W}_+^* \times Y_+^*)^* = W_+ \times Y_+$. 另一方面，$\varphi \in N(S^*)^\perp = R(S)$，故 $\exists x \in X$，使 $\varphi = Sx = (Tx, Ax)$. 于是 $Tx = w \leqslant 0, Ax = y \leqslant 0$. 由 $\langle \mathring{W}_+^* \times Y_+^*, \varphi \rangle = \langle \mathring{W}_+^*, w \rangle + \langle Y_+^*, y \rangle < 0$ 推出 $w \neq 0$，因此 $Tx < 0$，故 x 是(P)的解. □

参考文献：[72, 93, 95, 104, 145, 203].

§ 3.5 Minimax 定理

给定函数 $\varphi: D \times K \to \mathbf{R}$，令

$$\alpha = \inf_{x \in D} \sup \varphi(x, K), \beta = \sup_{y \in K} \inf \varphi(D, y). \tag{1}$$

φ 关联着一对极值问题：

$$\min \sup \varphi(x, K) = \alpha, \quad x \in D \tag{P}$$

与

$$\max \inf \varphi(D, y) = \beta, \quad y \in K. \tag{P*}$$

直接看出恒有 $\beta \leqslant \alpha$. 基本的问题是，在什么条件下 $\alpha = \beta$? 使 $\alpha = \beta$ 成立的三种特殊情况是

$$\min_{x \in D} \sup \varphi(x, K) = \max_{y \in K} \inf \varphi(D, y); \tag{2}$$

$$\min_{x \in D} \sup \varphi(x, K) = \sup_{y \in K} \inf \varphi(D, y); \tag{3}$$

$$\inf_{x \in D} \sup \varphi(x, K) = \max_{y \in K} \inf \varphi(D, y). \tag{4}$$

注意 (3) $\Rightarrow \alpha = \beta > -\infty$，(4) $\Rightarrow \alpha = \beta < \infty$. (2) 成立的充要条件是存在 $(\bar{x}, \bar{y}) \in D \times K$ 满足

$$\varphi(\bar{x},K) \leqslant \varphi(\bar{x},\bar{y}) \leqslant \varphi(D,\bar{y}). \tag{5}$$

当式(5)成立时称 $(\bar{x},\bar{y})$ 为 φ 的**鞍点**. 鞍点是极不常见的(参看§5.1). 式(3) 成立 $\Leftrightarrow$ 问题(P) 有解且 $\alpha = \beta$;式(4) 成立 $\Leftrightarrow$ 问题(P^*) 有解且 $\alpha = \beta$. 注意式(3) 与式(4) 互相对偶. 凡断言 $\alpha = \beta$ 或等式(2) ~ (4) 之一成立的结果称为 **minimax 定理**. 这类结果颇多,且对于最优化问题有重要应用.

令 $E = \mathbf{R}^D, F = \mathbf{R}^K$,二者依通常的积拓扑成为 LCS. 每个 $r \in \mathbf{R}$ 可看作 D 上恒等于 r 的函数,因此有自然的嵌入 $\mathbf{R} \subset E$. 同样认定自然的嵌入 $\mathbf{R} \subset F$. E 与 F 分别由闭凸锥 $E_+ = \mathbf{R}_+^D$ 与 $F_+ = \mathbf{R}_+^K$ 导入序 $\leqslant$. 对于 $u \in E$,令 $\operatorname{supp} u = \{x \in D : u(x) \neq 0\}$. 可验证 E 的拓扑对偶为

$$E^* = \{\rho \in E : \operatorname{supp} \rho \text{ 为有限集}\},$$

$\rho \in E^*$ 与 $u \in E$ 的配对为 $\langle \rho, u \rangle = \sum_x \rho(x) u(x)$. 显然 $E_+^* = E_+ \cap E^*$. F^* 与 F_+^* 仿此.

空间 E,F 的价值在于,对 φ 的研究可转化为考察由 φ 诱导的一对映射:

$$\begin{cases} \Phi : D \to F, x \to \varphi(x, \cdot), \\ \Psi : K \to E, y \to \varphi(\cdot, y), \end{cases} \tag{6}$$

从而有可能直接应用前几节的结果.

现在约定一些在本节起重要作用的记号. 字母 ω 或 ω' 恒记非空有限集. $\forall \omega \subset D, \omega' \subset K, \varepsilon > 0, r \in \mathbf{R}$,约定

$$\begin{cases} U(\omega,\varepsilon) = \{u \in E : |u(x)| \leqslant \varepsilon (\forall x \in \omega)\}, \\ V(\omega',\varepsilon) = \{v \in F : |v(x)| \leqslant \varepsilon (\forall y \in \omega')\}; \end{cases} \tag{7}$$

$$\begin{cases} D(\omega',r) = \{x \in D : \varphi(x,\omega') \leqslant r\}, \\ K(\omega,r) = \{y \in K : \varphi(\omega,y) \geqslant r\}; \end{cases} \tag{8}$$

$$\begin{cases} M_r = \Psi(K) - E_+ - r, M = M_0, \\ N_r = \Phi(D) + F_+ - r, N = N_0. \end{cases} \tag{9}$$

所有形如 $U(\omega,\varepsilon)$ 的集构成 E 的一个闭 0- 邻域基;关于 F 仿此.

大多数关于 φ 的 minimax 定理要用到某种凸性,这些凸性通

过映射 Φ,Ψ 表达最为简便.

3.5.1 定义 (i) $\varphi(x,y)$ 对 x 类凸、近类凸与**几乎类凸**分别意指 $\Phi: D \to F$ 为 F_+-类凸、F_+-近类凸与 F_+-次类凸;$\varphi(x,y)$ 对 y 类凹、近类凹与**几乎类凹**分别意指 $\Psi: K \to E$ 为 E_+-类凹、E_+-近类凹与 E_+-次类凹.

(ii) 若 $\forall t \in (0,1), \forall \varepsilon > 0, \forall x_1, x_2 \in D, \exists x \in D$,使

$$\Phi(x) \leqslant t\Phi(x_1) + t'\Phi(x_2) + \varepsilon,$$

则说 φ 对 x 次类凸;**近次类凸**意味着将次类凸条件中的"$\forall t \in (0,1)$"改为"$\exists t \in (0,1)$". φ 对 y 次类凹与近次类凹的意义仿此.

(iii) 若 φ 对 x 类凸而对 y 类凹,则说 φ 为**类鞍**、**近类鞍**、**次类鞍**等仿此.

注 1°"几乎类凸"一词是参照[82]中的"几乎类凹"概念而采用的.由 3.5.1,φ 对 x 几乎类凸等价于 Φ 为 F_+-次类凸,而"φ 对 x 次类凸"强于"φ 对 x 几乎类凸".就 φ 对 x 而言,一般地有:

$$\text{类凸} \Rightarrow \text{次类凸} \Rightarrow \text{近次类凸} \Rightarrow \text{几乎类凸}.$$

2°结合 3.5.1,3.2.6 与 3.2.2 得出,φ 对 x 几乎类凸 $\Leftrightarrow N = \Phi(D) + F_+$ 次凸 $\Leftrightarrow \exists t \in (0,1), \forall \omega \subset K, \forall \varepsilon > 0: tN + t'N \subset N + V(\omega,\varepsilon)$,后者相当于

$$\begin{cases} \exists t \in (0,1), \forall \omega \subset K, \forall \varepsilon > 0, \forall x_1, x_2 \in D, \exists x \in D, \\ \forall y \in \omega: \varphi(x,y) \leqslant t\varphi(x,y) + t'\varphi(x_2,y) + \varepsilon. \end{cases} \tag{10}$$

[82]中对"几乎类凹"有一类似于式(10)的直接刻画.

结合 3.5.1 与 3.2.6 直接得出以下命题.

3.5.2 命题 φ 对 x 类凸、近类凸与几乎类凸分别等价于 N 凸、近凸与 $\overline{N}$ 凸;φ 对 y 类凹、近类凹与几乎类凹分别等价于 M 凸、近凸与 $\overline{M}$ 凸,M,N 依式(9).

注 容易看出,φ 对 x 次类凸 $\Leftrightarrow N$ 满足:

$$\forall t \in (0,1), \forall \varepsilon > 0: tN + t'N \subset N + [-\varepsilon, \varepsilon].$$

此条件较"N 次凸"为强.

3.5.3 引理 设 $r \in \mathbf{R}$. 假定以下条件之一满足：(i) D 为 Hausdorff 空间，φ 对 x 为 lsc，$\exists \omega \subset K, \sigma > r: D(\omega,\sigma)$ 紧；(ii) X 为自反 B-空间，$D \subset X$ 弱闭，φ 对 x 为 wlsc，$\exists \lambda \in F_+^*: \sum_y \lambda(y) = 1$，且

$$\liminf_{x \in D, |x| \to \infty} \langle \lambda, \varphi(x, \cdot) \rangle > r.$$

则存在 F 的 0-邻域 V，使 $N_r \cap V$ 为闭集.

证 设 V 是 F 的闭 0-邻域，$\{v_i\}$ 是 $N_r \cap V$ 中的网，$v_i \to v$，则 $v \in V$. 取 $x_i \in D$，使 $\varphi(x_i, \cdot) \leqslant v_i + r$. 为证 $v \in N_r$，分两种情况考虑.

首先设条件(i)满足. 令 $\varepsilon = \sigma - r, V = V(\omega,\varepsilon)$，则 $\forall y \in \omega: \varphi(x_i, y) \leqslant v_i(y) + r \leqslant \varepsilon + r = \sigma$，可见 $x_i \in D(\omega,\sigma)$. 不妨设 $x_i \to x \in D$，则

$$\varphi(x,y) \leqslant \liminf_i \varphi(x_i,y) \leqslant v(y) + r \ (\forall y \in K),$$

这表明 $v \in N_r$.

其次设条件(ii)满足，取 $\varepsilon > 0$，使当 $|x| \to \infty$ 时，

$$\langle \lambda, \varphi(x, \cdot) \rangle > r + \varepsilon,$$

λ 依条件(ii). 取 V 为 F 的这样的闭 0-邻域：$\lambda(V) \subset [-\varepsilon,\varepsilon]$. 因 $\langle \lambda, \varphi(x_i, \cdot) \rangle \leqslant \langle \lambda, v_i \rangle + r \leqslant \varepsilon + r$，故 $\{x_i\}$ 有界. 不妨设 $x_i \rightharpoonup x \in D$，则同样有 $\varphi(x,y) \leqslant v(y) + r (\forall y \in K)$，于是 $v \in N_r$. □

现在应用 3.3.5（并结合 3.5.3）得出一个关于 φ 的"Gale 型定理"，它在本节中将起基本作用.

3.5.4 定理 假定：(i) φ 对 x 几乎类凸；(ii) 存在 F 的 0 邻域 V，使 $N \cap V$ 为闭集. 则以下命题两择一：

(P) $\exists x \in D: \varphi(x,K) \leqslant 0 (\Leftrightarrow \Phi(x) \leqslant 0)$；

(Q) $\exists \lambda \in F_+^*, \forall x \in D: \langle \lambda, \varphi(x, \cdot) \rangle > 0$.

条件(ii)可代以如下条件之一：(iii) D 为 Hausdorff 空间，φ 对 x 为 lsc，$\exists \omega \subset K, \sigma > 0: D(\omega,\sigma)$ 紧；(iv) X 为自反 B-空间，$D \subset X$ 弱闭，φ 对 x 为 wlsc，$\exists \lambda \in F_+^*$，使

$$\liminf_{x \in D, |x| \to \infty} \langle \lambda, \varphi(x, \cdot) \rangle > 0.$$

证 注意到定理条件(i),(ii)相当于 $\Phi: D \to F$ 满足 3.3.5 中的条件(i),(ii),而由 3.5.3 知(iii) $\Rightarrow$ (ii),(iv) $\Rightarrow$ (ii),由此立得定理结论. □

3.5.4 推广了[113,Th. 3.5]. 一个对偶的结果如下.

3.5.5 定理 假定:(i) φ 对 y 几乎类凹;(ii) 存在 E 的 0-邻域 U,使 $M \cap U$ 为闭集. 则以下命题两择一:

(P) $\exists y \in K: \varphi(D, y) \geqslant 0 (\Leftrightarrow \Psi(y) \geqslant 0)$;

(Q) $\exists \rho \in E_+^*, \forall y \in K: \langle \rho, \varphi(\cdot, y) \rangle < 0$.

条件(ii)可代以如下条件之一:(iii) K 为 Hausdorff 拓扑空间且 φ 对 y 为 usc;$\exists \omega \subset D, \sigma < 0: K(\omega, \sigma)$ 紧;(iv) Y 为自反 B-空间,$K \subset Y$ 弱闭,φ 对 y 为 wusc;$\exists \rho \in E_+^*$,使

$$\limsup_{y \in K, |y| \to \infty} \langle \rho, \varphi(\cdot, y) \rangle < 0.$$

下面利用 3.5.4 导出 minimax 定理. 首先建立

3.5.6 引理 设 φ 对 x 几乎类凸而对 y 次类凹;存在 $r < \alpha$ 与 F 的 0-邻域 V,使 $N_r \cap V$ 闭,则 $\forall \epsilon > 0, \exists y_\epsilon \in K$,使

$$r \leqslant \varphi(D, y_\epsilon) + \epsilon.$$

证 由 $r < \alpha$ 推出问题"$\exists x \in D: \varphi(x, K) \leqslant r$"无解. 以 $\varphi - r$ 代 φ,应用 3.5.4 得 $\lambda \in F_+^*: \langle \lambda, \varphi(x, \cdot) \rangle > \langle \lambda, r \rangle (\forall x \in D)$. 不妨设 $\sum_y \lambda(y) = 1$,因此有

$$\sum_y \lambda(y) \varphi(x, y) > r \ (\forall x \in D). \tag{11}$$

取定 $\epsilon > 0$. 因 φ 对 y 次类凹,用一归纳论证不难得出 $y_\epsilon \in K$,使得

$$\varphi(x, y_\epsilon) \geqslant \sum_y \lambda(y) \varphi(x, y) - \epsilon \ (\forall x \in D). \tag{12}$$

结合(11)与(12)得出 $r < \varphi(x, y_\epsilon) + \epsilon \ (\forall x \in D)$,即得所要证. □

3.5.7 定理 假定:(i) φ 对 x 几乎类凸而对 y 次类凹;(ii) 存

在任意接近 β 的 $r > \beta$ 与 F 的 0-邻域 V，使 $N_r \cap V$ 闭，则 $\alpha = \beta < \infty$. 若条件(ii) 代以如下条件之一：(iii) D 为 Hausdorff 空间，φ 对 x 为 lsc，$\exists \omega \subset K, \sigma > \beta$；$D(\omega,\sigma)$ 紧；(iv) X 为自反 B-空间，$D \subset X$ 弱闭，φ 对 x 为 wlsc，$\exists \lambda \in F_+^*$，使 $\sum_y \lambda(y) = 1$，且

$$\liminf_{x \in D, |x| \to \infty} \langle \lambda, \varphi(x, \cdot) \rangle > \beta,$$

则(3)成立且 $\alpha = \beta \neq \pm \infty$.

证 若 $\alpha \neq \beta$，则 $\beta < \alpha$. 由条件(ii) 有 $r \in (\beta,\alpha)$，F 的 0-邻域 V，使 $N_r \cap V$ 闭. $\forall \varepsilon > 0$，由 3.5.6 有 $y_\varepsilon \in K$，使

$$r \leqslant \inf \varphi(D, y_\varepsilon) \leqslant \beta + \varepsilon.$$

令 $\varepsilon \downarrow 0$ 得 $r \leqslant \beta$，这矛盾于 $r > \beta$. 因此 $\alpha = \beta < \infty$.

设条件(iii) 满足. $\forall r \in (\beta,\sigma)$，由 3.5.3 知有 F 的 0-邻域 V，使 $N_r \cap V$ 闭. 于是由已证结论有 $\alpha = \beta < \infty$. $\forall n \geqslant 1$，取 $x_n \in D$，使 $\sup \varphi(x_n, K) \downarrow \alpha$. 显然 $x_n \in D(\omega,\sigma)$，故不妨设 $x_n \to \bar{x} \in D$. 由

$$\varphi(\bar{x}, y) \leqslant \liminf_n \varphi(x_n, y) \leqslant \alpha \ (\forall y \in K)$$

得 $\alpha = \sup \varphi(\bar{x}, K)$，因此(3) 成立且 $\alpha = \beta \neq \pm \infty$.

当条件(iv) 满足时，类似地可证上述结论. □

对偶地有下面定理.

3.5.8 定理 假定：(i) φ 对 x 次类凸而对 y 几乎类凹；(ii) 存在任意接近于 α 的 $r < \alpha$ 与 E 的 0-邻域 U，使 $M_r \cap U$ 闭，则 $\alpha = \beta > -\infty$. 若条件(ii) 代以如下条件之一：(iii) K 为 Hausdorff 空间，φ 对 y 为 usc，$\exists \omega \subset D, \sigma < \alpha$；$K(\omega,\sigma)$ 紧；(iv) Y 为自反 B-空间，$K \subset Y$ 弱闭，φ 对 y 为 wusc，$\exists \rho \in E_+^*$，使 $\sum_x \rho(x) = 1$，且

$$\limsup_{y \in K, |y| \to \infty} \langle \rho, \varphi(\cdot, y) \rangle < \alpha,$$

则(4)成立且 $\alpha = \beta \neq \pm \infty$.

若 D 或 K 本身紧，则 3.5.7 中的 $D(\omega,\sigma)$ 或 3.5.8 中的 $K(\omega,\sigma)$ 自动为紧集. 许多 minimax 定理正是直接以 D 或 K 紧为条件(参看[24,105,159,200]). 下面不加证明地引述这类结果中的主

要者. 约定以下术语：若 φ 对 x 拟凸而对 y 拟凹，则说 φ 为**拟鞍**；若 φ 对 x 为 lsc 而对 y 为 usc，则说 φ 为**半连续**.

3.5.9 定理 以下每个条件推出(3)成立(因而 $\alpha = \beta$)：

(i) D 为紧 Hausdorff 空间，φ 近次类鞍且对 x 为 lsc；

(ii) $D \subset X$ 凸紧，$K \subset Y$ 凸，φ 拟鞍且半连续；

(iii) $D \subset X$ 凸紧，φ 对 x 拟凸且 lsc，对 y 近次类凹.

以下是对偶的结果.

3.5.10 定理 下面每个条件推出(4)成立(因而 $\alpha = \beta$)：

(i) K 为紧 Hausdorff 空间，φ 近次类鞍且对 y 为 usc；

(ii) $D \subset X$ 凸，$K \subset Y$ 凸紧，φ 拟鞍且半连续；

(iii) $K \subset Y$ 凸紧，φ 对 x 近次类凸，对 y 拟凹且 usc.

注 3.5.9 与 3.5.10 都不能断定 $\alpha = \beta \neq \pm\infty$. 例如，$\varphi(x, y) = x - y^2$ 在 $\mathbf{R} \times [-1,1]$ 上是连续鞍函数，3.5.10 中条件(i)～(iii) 皆满足，但

$$\inf_{x \in \mathbf{R}} \sup_{y \in [-1,1]} (x - y^2) = \max_{y \in [-1,1]} \inf_{x \in \mathbf{R}} (x - y^2) = -\infty.$$

设 $f : D \to \mathbf{R}$，令

$$C = \{(r,v) \in \mathbf{R} \times F : f(x) \leqslant r, \varphi(x, \cdot) \leqslant v (\exists x \in D)\}. \tag{13}$$

注意到 $C = (f,\Phi)(D)$，Φ 依(6)、不难从 3.4.10 直接推出如下"Motzkin 型定理".

3.5.11 定理 假定：(i) (f,Φ) 为 $\mathbf{R}_+ \times F_+$ - 次类凸；(ii) 存在 $\sigma \geqslant 0$ 与 F 的 0 邻域 V，使 $C \cap ((-\infty,\sigma] \times V)$ 为闭集. 则以下命题两择一：

(P) $\exists x \in D$：$f(x) < 0, \varphi(x,K) \leqslant 0$；

(Q) $\forall \varepsilon > 0, \exists (\rho,\lambda) \in \mathbf{R}_+ \times F_+^*, \forall x \in D$，有

$$\rho f(x) + \langle \lambda, \varphi(x, \cdot)\rangle > -\rho\varepsilon.$$

若假定：(iii) $\exists \hat{x} \in D$：$\varphi(\hat{x},K) \leqslant 0$，则(Q)可写成：

$$\forall \varepsilon > 0, \exists \lambda \in F_+^*, \forall x \in D: f(x) + \langle \lambda, \varphi(x, \cdot)\rangle > -\varepsilon.$$

条件(ii)可代以如下条件之一：(iv) D 为 Hausdorff 空间，f 与 φ 对

x 为 lsc；$\exists \omega \subset K, \sigma \geqslant 0, \varepsilon > 0$，使 $\{x \in D : f(x) \leqslant \sigma, \varphi(x,\omega) \leqslant \varepsilon\}$ 为紧集；(v) X 为自反 B- 空间，$D \subset X$ 弱闭，f 与 φ 对 x 为 wlsc，$\exists (\rho,\lambda) \in \mathbf{R}_+ \times F_+^*$，使

$$\liminf_{x \in D, |x| \to \infty} \rho f(x) + \langle \lambda, \varphi(x, \cdot) \rangle > 0.$$

注 1° (f,Φ) 为 $\mathbf{R}_+ \times F_+$ - 次类凸意味着 C 次凸，即

$$\exists t \in (0,1), \forall \varepsilon > 0, \forall \omega \subset K:$$
$$tC + t'C \subset C + [-\varepsilon,\varepsilon] \times V(\omega,\varepsilon).$$

这又相当于

$$\begin{cases} \exists t \in (0,1), \forall \varepsilon > 0, \forall \omega \subset K, \forall x_1, x_2 \in D, \exists x \in D: \\ f(x) \leqslant t f(x_1) + t' f(x_2) + \varepsilon, \\ \varphi(x,y) \leqslant t\varphi(x_1,y) + t'\varphi(x_2,y) + \varepsilon \ (\forall y \in \omega). \end{cases} \tag{14}$$

2° 3.5.11 是[113, Th. 3.1]的一个推广. 交换 x, y，一个类似的对偶结果可直接写出.

本节的结果可表为"无限不等式组"的形式，只要将 K(它通常是无限集) 看作指标集，相应地令 $f_k(x) = \varphi(x,k), (x,k) \in D \times K$. 函数组 $\{f_k : k \in K\}$ 类凸意味着 φ 对 x 类凸；次类凸等仿此. 任给 $v \in F$，将 v 表为 $(v_k)_{k \in K}$. 作为例子，下面将 3.5.4 改写成：

3.5.12 定理 假定：(i) $\{f_k\}$ 几乎类凸；(ii) 存在 F 的 0- 邻域 V，使 $N \cap V$ 闭，其中 $N = \{v \in F : f_k(x) \leqslant v_k (\exists x \in D, \forall k \in K)\}$. 则以下命题两择一：

(P) $\exists x \in D, \forall k \in K : f_k(x) \leqslant 0$；

(Q) $\exists \lambda \in F_+^*, \forall x \in D : \sum_k \lambda_k f_k(x) > 0.$

条件(ii) 可代以如下条件之一：(iii) D 是 Hausdorff 空间，每个 f_k 为 lsc；$\exists \omega \subset K, \varepsilon > 0 : \{x \in D : f_k(x) \leqslant \varepsilon (\forall k \in \omega)\}$ 为紧集；(iv) X 是自反 B- 空间，$D \subset X$ 弱闭，每个 f_k 为 wlsc；

$$\exists \lambda \in F_+^* : \liminf_{x \in D, |x| \to \infty} \sum_k \lambda_k f_k(x) > 0.$$

在［194］中可看到类似的结果.

参考文献:［24,43,82,93,105,113,119,151,159,169,172,192,194,200,202］.

§3.6 Minimax 定理导出的择一定理

设 $\varphi(x,y):D\times C\to\mathbf{R}$,$C$ 是某个 LCS 中的锥. 若 $0\overline{\in}K\subset C=\mathbf{R}_+K$,则称 K 为锥 C 的基. 若 $\delta>0$,

$$\varphi(x,ty)=t^{\delta}\varphi(x,y)(\forall x\in D,y\in K,t\geqslant 0),$$

则说 φ 对 y 是 δ 次正齐次的. 保留上节中与 φ 有关的记号.

上节的等式(4)蕴涵以下命题:

$$\inf_{x\in D}\sup\varphi(x,K)\geqslant 0\Leftrightarrow\max_{y\in K}\inf\varphi(D,y)\geqslant 0. \tag{1}$$

基于(1)直接得到:

3.6.1 引理 设 C 有基 K,φ 对 y 为 δ 次正齐次且满足式(1),则以下命题两择一:

(P) $\exists x\in D:\varphi(x,C\backslash\{0\})<0$;

(Q) $\exists y\in C\backslash\{0\}:\varphi(D,y)\geqslant 0$.

结合 3.6.1 与 3.5.8,3.5.10 得到:

3.6.2 定理 设 C 有基 K,φ 对 y 为 δ 次正齐次且 usc,则以下每个条件推出 3.6.1 的结论:

(i) φ 对 x 次类凸而对 y 在 K 上几乎类凹,$\exists\omega\subset D,\sigma<0$,使 $K(\omega,\sigma)$ 紧;

(ii) Y 为自反 B-空间,$K\subset Y$ 弱闭,φ 对 x 次类凸、对 y 在 K 上几乎类凹且 wusc,

$$\exists x\in D:\limsup_{y\in K,\|y\|\to\infty}\varphi(x,y)<0; \tag{2}$$

(iii) K 紧且 φ 在 $D\times K$ 上近次类鞍;

(iv) $D\subset X$ 凸,K 凸紧,φ 在 $D\times K$ 上拟鞍且半连续;

(v) K 凸紧，φ 对 x 近次类凸且对 y 在 K 上拟凹.

3.6.2 综合并推广了[43,82]中的一些结果.

应用 3.6.2 的最简单的方式是取 $\varphi(x,\lambda) = \lambda g(x)$，$C = Y_+^*$. 直接看出：$\varphi$ 对 λ 是1次齐次的；关于 Y^* 中的弱 * 拓扑 φ 对 λ 连续；$K \subset Y_+^*$ 凸 $\Rightarrow$ φ 对 λ 在 K 上凹；若 g 为 Y_+-次类凸且 K 紧，则 $\varphi|(D \times K)$ 对 x 次类凸. 于是从 3.6.2 推出：

3.6.3 推论 设 Y_+^* 有弱 * 紧凸基 K，且以下条件之一满足：(i) g 为 Y_+-次类凸；(ii) g 为 $*$ 拟凸且 $*$lsc. 则以下命题两择一：

(P) $\exists x \in D: \langle Y_+^* \backslash \{0\}, g(x)\rangle < 0$;

(Q) $\exists \lambda \in Y_+^* \backslash \{0\}: \langle \lambda, g(D)\rangle \geqslant 0$.

3.6.3 可看作一个"Gordan 型定理"而与 3.3.2 相对照. 若 $Y_+^{\circ} \neq \varnothing$，则 $\langle Y_+^* \backslash \{0\}, g(x)\rangle < 0 \Leftrightarrow g(x) \ll 0$(1.4.4)，此时 3.6.3 与 3.3.2 的结论相同，而条件不尽相同. 任取 $y \in Y_+^{\circ}$，可验证 $K \overset{\text{def}}{=} \{\lambda \in Y_+^* : \lambda(y) = 1\}$ 是 Y_+^* 的弱 * 紧凸基([114]). 于是利用 3.6.3 得到 3.3.2 的如下补充.

3.6.4 推论 若 $Y_+^{\circ} \neq \varnothing$，$g$ 为 $*$ 拟凸且 $*$lsc，则 3.3.2 的结论成立.

在 3.6.3 中取 $g = A \in L(X,Y)$ 得：

3.6.5 推论 设 Y_+^* 有弱 * 紧凸基，$D \subset X$ 凸，则以下命题两择一：

(P) $\exists x \in D: \langle Y_+^* \backslash \{0\}, Ax\rangle < 0$;

(Q) $\exists \lambda \in Y_+^* \backslash \{0\}: A^*\lambda \in D^*$.

若 $Y_+^{\circ} \neq \varnothing$，则 3.6.5 的结论已包含于 3.3.3.

今后常用到如下的弱 **Slater 条件**：

$$\exists \hat{x} \in D: \langle Y_+^* \backslash \{0\}, g(\hat{x})\rangle < 0. \tag{3}$$

若 $Y_+^{\circ} \neq \varnothing$，则(3) $\Leftrightarrow$ $\exists \hat{x} \in D: g(\hat{x}) \ll 0$(所谓 **Slater 条件**).

以 (f,g) 代 g 从 3.6.3 得到：

3.6.6 推论 设 W_+^* 与 Y_+^* 分别有弱 * 紧凸基 K_1 与 K_2；(f,g)

为 $W_+ \times Y_+$ - 次类凸，或为 * 拟凸且 * lsc；条件(3) 满足. 则以下命题两择一(参考[43])：

(P) $\exists x \in D: \langle W_+^* \setminus \{0\}, f(x) \rangle < 0, g(x) \leqslant 0$;

(Q) $\exists (\rho, \lambda) \in (W_+^* \setminus \{0\}) \times Y_+^*: (\rho f + \lambda g)(D) \geqslant 0$.

证 令 $K = \{(t\rho, t'\lambda): t \in J, \rho \in K_1, \lambda \in K_2\}$，则可验明 K 是 $W_+^* \times Y_+^*$ 的弱*紧凸基. 若问题

$$\exists x \in D: \langle (W_+^* \times Y_+^*) \setminus \{0\}, (f(x), g(x)) \rangle < 0 \qquad (\mathrm{P}')$$

有解 x，则易验知 x 亦是(P) 的解. 因此，若(P) 无解，则(P′) 亦无解. 于是由 3.6.3 有 $0 \neq (\rho, \lambda) \in W_+^* \times Y_+^*: (\rho f + \lambda g)(D) \geqslant 0$. 必定 $\rho \neq 0$，否则由条件(3) 将有 $0 \leqslant \lambda g(\hat{x}) < 0$! 因此(Q) 有解. □

注意条件(3) 仅用于以上证明的最后一步，即用于说明 $\rho \neq 0$. (3) 亦可代以条件 $Y_+^* \cap g(D)^* = \{0\}$.

3.6.6 的线性形式是下述推论.

3.6.7 推论 设 W_+^* 与 Y_+^* 均有弱*紧凸基，$D \subset X$ 凸；$\exists \hat{x} \in D: \langle Y_+^* \setminus \{0\}, A\hat{x} \rangle < 0$(或 $Y_+^* \cap (AD)^* = \{0\}$)，则以下命题两择一：

(P) $\exists x \in D: \langle W_+^* \setminus \{0\}, Tx \rangle < 0, Ax \leqslant 0$;

(Q) $\exists (\rho, \lambda) \in (W_+^* \setminus \{0\}) \times Y_+^*: T^*\rho + A^*\lambda \in D^*$.

3.6.6 与 3.6.7 一起推广了 3.4.8. 若 $W = \mathbf{R}^n, W_+ = \mathbf{R}^n_+$，则不必提到"$W_+^*$ 有弱*紧凸基"；若 $W = \mathbf{R}, W_+ = \mathbf{R}_+$，则 3.6.6 与 3.6.7 中的命题(Q) 可分别改换为"$\exists \lambda \in Y_+^*: (f + \lambda g)(D) \geqslant 0$"与"$\exists \lambda \in Y_+^*: T + A^*\bar{\lambda} \in D^*$"，后者又等价于 $T \in A^*Y^* + D^*$.

再进一步，以 (f, g, h) 代 g 类似于 3.6.3 有下面推论.

3.6.8 推论 设 W_+^*，Y_+^* 与 Z^* 分别有弱*紧基 K_1, K_2 与 K_3，$K = \{(t_1u_1, t_2u_2, t_3u_3): t_i \geqslant 0, \sum t_i = 1, u_i \in K_i (1 \leqslant i \leqslant 3)\}$，$L(\cdot, \rho, \lambda, \mu) = \rho f + \lambda g + \mu h$. 若 L 在 $D \times K$ 上近次类鞍，则以下命题至少一个成立：

(P) $\exists x \in D: \langle W_+^* \setminus \{0\}, f(x) \rangle < 0, g(x) \leqslant 0, h(x) = 0$;

(Q) $\exists(\rho,\lambda,\mu)\in(W_+^*\times Y_+^*\times Z^*)\backslash\{0\}:L(D,\rho,\lambda,\mu)\geqslant 0$.

若附设条件:$\exists\hat{x}\in D$: $\langle Y_+^*\backslash\{0\},g(\hat{x})\rangle<0,h(\hat{x})=0$; $h(D)^*=\{0\}$,则可设(Q)中的 $\rho\neq 0$;当 $W=\mathbf{R}$ 时可设 $\rho=1$. 此时(P)与(Q)两择一.

证明类似于3.6.6,只需说明:K 是 $W_+^*\times Y_+^*\times Z^*$ 的弱*紧基;命题(P)蕴涵于命题

$\exists x\in D:\langle(W_+^*\times Y_+^*\times Z^*)\backslash\{0\},(f(x),g(x),h(x))\rangle<0$.

3.6.8可作为一个"Motzkin型定理"而与3.4.2相对照.当 $W_-^\circ\neq\varnothing$ 时,3.6.8的结论与3.4.2一致,而条件有所不同.3.6.8的条件不甚理想.首先,要求 Z^* 有弱*紧基是一强的限制(除非 $\dim Z<\infty$).其次,K 一般非凸,要验证 L 对 (ρ,λ,μ) 在 K 上近次类凹颇成问题.鉴于此,3.6.8不及3.6.3~3.6.7有用.

与§3.3~§3.4的结果比较,3.6.3~3.6.8的一个突出优点是避开了条件 $W_+^\circ\neq\varnothing\neq Y_+^\circ$.在无限维空间中,这是很有意义的.

若对 $\varphi(x,\lambda)=\lambda g(x)$ 不是应用3.6.2,而是应用3.5.4,则有与3.6.3不同的如下择一定理.

3.6.9 定理 设 Y_+^* 有凸基 K,且以下条件之一满足:(i) 存在 $E=\mathbf{R}^D$ 的0-邻域 U,使 $\widetilde{U}=\{u\in U:u\leqslant\lambda g(\exists\lambda\in K)\}$ 在 E 中闭;(ii) $\exists\omega\subset D,\sigma<0$,使

$$K(\omega,\sigma)\overset{\text{def}}{=}\{\lambda\in K:\lambda g(x)\geqslant\sigma(\forall x\in\omega)\}\tag{4}$$

为弱*紧集;(iii) Y 为自反 B-空间,K 弱闭,且

$$\exists\rho\in E_+^*:\limsup_{\lambda\in K,\|\lambda\|\to\infty}\langle\lambda,\sum_x\rho(x)g(x)\rangle<0,\tag{5}$$

则以下命题两择一:

(P) $\exists\rho\in E_+^*:\langle Y_+^*\backslash\{0\},\sum_x\rho(x)g(x)\rangle<0$;

(Q) $\exists\lambda\in Y_+^*\backslash\{0\}:\langle\lambda,g(D)\rangle\geqslant 0$.

所宜注意者,3.6.9对于 g 没有任何凸性要求.3.6.9中的(Q)与3.6.3中的(Q)一致;若 $\sup\rho=\{x\}$,或 g 为 Y_+-次类凸,则3.

6.9中的(P)与3.6.3中的(P)一致.注意(5)可代以较易验证的条件(参照(2)):

$$\exists x\in D:\limsup_{\lambda\in K,\,\lambda\to\infty}\lambda g(x)<0. \tag{6}$$

对 $\varphi(x,\lambda)=\lambda g(x)$ 应用3.5.8得出:

3.6.10 定理 设 Y_+^* 有凸基 K,g 为 Y_+-类凸,且以下条件之一满足:(i) $\exists\omega\subset D,\sigma<\alpha\overset{\text{def}}{=}\inf_{x\in D}\sup_{\lambda\in K}\lambda g(x)$,使 $K(\omega,\sigma)$(依(4))为弱*紧;(ii) Y 为自反 B 空间,K 弱闭,且 $\exists\rho\in E_+^*$,使得 $\sum_x\rho(x)=1$,

$$\limsup_{\lambda\in K,\,|\lambda|\to\infty}\langle\lambda,\sum_x\rho(x)g(x)\rangle<\alpha. \tag{7}$$

则
$$\inf_{x\in D}\sup_{\lambda\in K}\lambda g(x)=\max_{\lambda\in K}\inf_{x\in D}\lambda g(x). \tag{8}$$

现在回到一般的函数 $\varphi:D\times K\to\mathbf{R}$. 令

$$\Lambda=\{\lambda\in F_+^*:\sum_y\lambda(y)=1\}, \tag{9}$$

则 Λ 是 F_+^* 的弱*闭凸基.分别以 F_+^*,Λ,Φ(依§3.5(6))替代 Y_+^*,K,g,从3.6.9与3.6.10得到下面的推论.

3.6.11 推论 设以下条件之一满足:(i) 存在 E 的0邻域 U,使 $\{u\in U:\sum_y\lambda(y)\varphi(\cdot,y)\geqslant u(\exists\lambda\in\Lambda)\}$ 为闭集;(ii) $\exists\omega\subset D,\sigma<0$,使

$$\Lambda(\omega,\sigma)\overset{\text{def}}{=}\{\lambda\in\Lambda:\langle\lambda,\varphi(x,\cdot)\rangle\geqslant\sigma(\forall x\in\omega)\} \tag{10}$$

为弱*紧集.则以下命题两择一:

(P) $\exists\rho\in E_+^*,\forall\lambda\in F_+^*\backslash\{0\}:\sum_{x,y}\rho(x)\lambda(y)\varphi(x,y)<0$;

(Q) $\exists\lambda\in F_+^*,\forall x\in D:\sum_y\lambda(y)\varphi(x,y)\geqslant0$.

注意3.6.11对 φ 没有任何凸性要求.

3.6.12 推论 设 φ 对 x 类凸,$\exists\omega\subset D,\sigma<\alpha$,其中

$$\alpha=\inf_{x\in D}\sup_{\lambda\in\Lambda}\langle\lambda,\varphi(x,\cdot)\rangle,$$

使 $\Lambda(\omega,\sigma)$(依(10))为弱*紧,则

$$\inf_{x\in D,\lambda\in\Lambda}\sup\sum_{y}\lambda(y)\varphi(x,y)=\max_{\lambda\in\Lambda,x\in D}\inf\sum_{y}\lambda(y)\varphi(x,y).$$

如同由(1)推出 3.6.1 中的两择一结论一样，由 3.6.12 立即得到：

3.6.13 推论 在 3.6.12 的条件下(假设其中 $\sigma<0$)，以下命题两择一：

(P) $\exists x\in D,\forall\lambda\in\Lambda$: $\sum_{y}\lambda(y)\varphi(x,y)<0$;

(Q) $\exists\lambda\in\Lambda,\forall x\in D$: $\sum_{y}\lambda(y)\varphi(x,y)\geqslant 0$.

与 3.6.11 对比，3.6.13 增加了一个凸性条件，因而(P)与(Q)呈现出更对称的形式.

若令 $f_k(x)=\varphi(x,k)$，则可将 3.6.13 改写成：

3.6.14 定理 设 $f_k:D\to\mathbf{R}$，$\{f_k:k\in K\}$ 类凸；$\exists\omega\subset D$，$\sigma<0$，使集

$$\left\{\lambda\in\Lambda:\sum_{k}\lambda_k f_k(x)\geqslant\sigma(\forall x\in\omega)\right\}$$

弱*紧，则以下命题两择一：

(P) $\exists x\in D,\forall\lambda\in\Lambda$: $\sum_{k}\lambda_k f_k(x)<0$;

(Q) $\exists\lambda\in\Lambda,\forall x\in D$: $\sum_{k}\lambda_k f_k(x)\geqslant 0$.

3.6.14 可与 3.5.12 对照.

参考文献：[43, 80, 82, 93, 114, 169, 194].

第四章　一阶最优性条件

本章考虑一般最优化问题

$$\min f(x),\ g(x)\leqslant 0,\ h(x)=0,\ x\in D \qquad (\mathrm{P})$$

的一阶最优性条件. 在问题(P)中,$f:X\to\mathbf{R}$ **是目标函数**,$g:X\to Y$ 与 $h:X\to Z$ **是约束函数**,$D\subset X$ **是约束集**,Y 中的序 $\leqslant$ 由闭凸锥 Y_+ 导入. 约定 $G=g^{-1}(Y_-)$, $H=h^{-1}(0)$, $M=D\cap G\cap H$. 称 M 为问题(P)的**可行集**,称每个 $x\in M$ 为(P)的**可行点**,通常以 $\bar{x}$ 记某个给定的可行点. 称 $f|M$ 的[局部]极小点为问题(P)的[局部]**最优解**;称 $f|M$ 的严格[局部]极小点为(P)的**严格[局部]最优解**. 最优解也称为最优点.

约定一些本章(及以下各章)常用的记号:$T=f'(\bar{x})$,$A=g'(\bar{x})$,$B=h'(\bar{x})$(只要这些导数存在);$K_g=Y_-+\mathbf{R}g(\bar{x})$, $Q=Y_+^*\cap\{g(\bar{x})\}^{\perp}=-K_g^*$. 若 $Y=\mathbf{R}^m$, 则 $Y_+=\mathbf{R}_+^m$, $g=(g_i)_{1\leqslant i\leqslant m}$, $I=\{i:g_i(\bar{x})=0\}$, $g_I=(g_i)_{i\in I}$, $A_I=g'_I(\bar{x})$,

$$\mathbf{R}^I=\{(y_i)\in\mathbf{R}^m: i\notin I\Rightarrow y_i=0\},$$

$$\mathbf{R}_+^I=\mathbf{R}^I\cap\mathbf{R}_+^m,\ \mathbf{R}_-^I=-\mathbf{R}_+^I.$$

定义问题(P)的 **Lagrange** 函数为

$$L(x,\rho,\lambda,\mu)=\rho f(x)+\lambda g(x)+\mu h(x),$$

其中 $(x,\rho,\lambda,\mu)\in D\times\mathbf{R}_+\times Y_+^*\times Z^*$. 当取 $\rho=1$ 或 $Z=\{0\}$ 时, 自动地将 $L(x,\rho,\lambda,\mu)$ 记作 $L(x,\lambda,\mu)$ 或 $L(x,\rho,\lambda)$;$L(x,\lambda)$ 与 $L(x,\mu)$ 的意义仿此. 本章的基本思想是,将问题(P)转化为关于函数 $L(\cdot,\rho,\lambda,\mu)$ 的无约束极值问题,因而可用 $L(\cdot,\rho,\lambda,\mu)$ 的一阶导数或次微分表达问题(P)存在最优解的必要或充分条件,这类条件称为**一阶最优性条件**. 分别以 L_x 与 $\partial_x L$ 记 L 对 x 的导数与次微分,则

$$L_x(\bar{x},\rho,\lambda,\mu) = T^*\rho + A^*\lambda + B^*\mu;$$
$$\partial_x L(\bar{x},\rho,\lambda,\mu) \subset \partial f_\rho(\bar{x}) + \partial g_\lambda(\bar{x}) + \partial h_\mu(\bar{x}),$$
其中 $f_\rho = \rho f, g_\lambda, h_\mu$ 仿此.

§4.1　可行集的切锥

设定 $\bar{x} \in M = G \cap H, G = g^{-1}(Y_-), H = h^{-1}(0)$. 为描述 $f|M$ 在 $\bar{x}$ 取极值的条件，经常要用到 M 在 $\bar{x}$ 的某种切锥（切锥的概念与记号依 §2.5). 在一定条件下，M 在 $\bar{x}$ 的切锥可通过约束函数在 $\bar{x}$ 的导数或次微分来刻画.

首先假定 g,h 在 $\bar{x}$ 可微.

4.1.1 定理　(i) $K_M(\bar{x}) \subset A^{-1}(Y \cap Q^\perp) \cap N(B) = \{x \in N(B) : \langle Q, Ax\rangle \leqslant 0\}$, $-K_M^*(\bar{x}) \supset A^*Q + R(B^*)$.

(ii) 设 $g,h \in C^1$, $(A,B): X \to Y \times Z$ 为满射，则 $K_M(\bar{x}) \supset (A^{-1}K_g) \cap N(B)$；当 K_g 闭时 $K_M(\bar{x}) = (A^{-1}K_g) \cap N(B)$；若再设 $A^*Q + R(B^*)$ 为弱*闭，则 $-K_M^*(\bar{x}) = A^*Q + R(B^*)$.

证　(i) 若 $z \in K_M(\bar{x})$，则 $\exists t_n \downarrow 0, z_n \to z$; $x_n \stackrel{\text{def}}{=} \bar{x} + t_n z_n \in M$(§2.5(1)). 由 $g(x_n) \leqslant 0$ 有
$$0 \geqslant \langle Q, g(x_n)\rangle = \langle Q, g(\bar{x}) + t_n A z_n + o(t_n)\rangle$$
$$= \langle Q, t_n A z_n + o(t_n)\rangle,$$
这得出 $\langle Q, Az\rangle \leqslant 0$. 类似地有 $Bz = 0$，于是 $z \in A^{-1}(Y \cap Q^*) \cap N(B) \stackrel{\text{def}}{=} K$. 故得 $K_M(\bar{x}) \subset K$. 进而有
$$-K_M^*(\bar{x}) \supset -K^* \supset (A^{-1}(Y \cap Q^*))^* + N(B)^\perp$$
$$\supset A^*(Y \cap Q^*)^* + R(B^*) \supset A^*Q + R(B^*).$$

(ii) 任取 $z \in (A^{-1}K_g) \cap N(B)$. 因 $\varphi \stackrel{\text{def}}{=} (g,h) \in C^1$, $\varphi'(\bar{x}) = (A,B)$ 是满射，故方程
$$\varphi(\bar{x} + y_n + n^{-1}z) = \varphi(\bar{x}) + n^{-1}\varphi'(\bar{x})z$$
有解 $y_n = o(1/n)$(1.2.6). 令 $z_n = ny_n + z$，则 $z_n \to z$, $x_n \stackrel{\text{def}}{=} \bar{x} +$

$n^{-1}z_n - \bar{x} + y_n + n^{-1}z$,

$$(g(x_n), h(x_n)) = \varphi(x_n) = \varphi(\bar{x}) + n^{-1}\varphi'(\bar{x})z$$
$$= (g(\bar{x}) + n^{-1}Az, 0).$$

由 $z \in A^{-1}K_g$ 知 $\exists r \in \mathbf{R}$,使 $Az \leqslant rg(\bar{x})$. 于是

$$g(x_n) = g(\bar{x}) + n^{-1}Az$$
$$\leqslant (1 + n^{-1}r)g(\bar{x}).$$

可见当 n 充分大时 $g(x_n) \leqslant 0$,而 $h(x_n) = 0$,故 $x_n \in M$. 因此 $z \in K_M(\bar{x})$、故得$(A^{-1}K_g) \cap N(B) \subset K_M(\bar{x})$. 若 K_g 闭,则 $K_g = Y \cap K_g^{**} = Y \cap Q^*_-$ (1.4.3),于是结合(i) 得 $K_M(\bar{x}) = (A^{-1}K_g) \cap N(B)$. 若 $\Phi \overset{\text{def}}{=} A^*Q + R(B^*)$ 弱* 闭,$u \in X^* \backslash \Phi$,则由 1.1.3 有 $z \in X$:$u(z) < \langle \Phi, z \rangle - \langle Q, Az \rangle + \langle R(B^*), z \rangle$. 由此得 $u(z) < 0$, $\langle Q, Az \rangle \geqslant 0$, $\langle R(B^*), z \rangle \geqslant 0$,从而 $z \in A^{-1}(Y \cap Q^*) \cap N(B) = -K_M(\bar{x})$,可见 $z \overline{\in} -K_M^*(\bar{x})$. 这与(i) 一起得出 $-K_M^*(\bar{x}) = \Phi = A^*Q + R(B^*)$. □

4.1.2 定理 (i) $K_G(\bar{x}) \subset A^{-1}(Y \cap Q^*_-)$, $K_G^*(\bar{x}) \supset -A^*Q$.

(ii) 设 $g \in C^1$, $R(A) + K_g = Y$,则 $K_G(\bar{x}) \supset A^{-1}K_g$, $K_G^*(\bar{x}) \subset (A^{-1}K_g)^* = -A^*Q$;当 K_g 为闭集时 $K_G(\bar{x}) = A^{-1}K_g$, $K_G^*(\bar{x}) = -A^*Q$.

证 (i) 取 $Z = \{0\}$ 从 4.1.1(i) 得出.

(ii) 关于 $K_G(\bar{x}) \supset A^{-1}K_g$ 参看[150]. 若 K_g 闭,则结合(i) 得 $K_G(\bar{x}) = A^{-1}K_g$. 余下只需证$(A^{-1}K_g)^* \subset -A^*Q$. 令 $T(x, r) = Ax + rg(\bar{x})$,则 $T \in L(X \times \mathbf{R}, Y)$, $R(T) + Y_- = R(A) + K_g = Y$,于是$(T^{-1}Y_-)^* = T^*Y^*_-$ (3.1.2). 若 $u \in (A^{-1}K_g)^*$,则$(u, 0) \in (T^{-1}Y_-)^*$. 于是 $\exists \lambda \in Y^*_-$:$(u, 0) = -T^*\lambda$. 由此易见 $\lambda \in Q$, $u = -A^*\lambda \in -A^*Q$. □

4.1.3 定理 $K_H(\bar{x}) \subset N(B)$, $K_H^*(\bar{x}) \supset R(B^*)$. 若 $h \in C^1$, $R(B) = Z$,则 $K_H(\bar{x}) = N(B)$, $K_H^*(\bar{x}) = R(B^*)$.

证 取 $Y = \{0\}$ 从 4.1.1 得出. □

一般无需用到可行锥 $F_M(\bar{x})$. 不过，关于 $F_G(\bar{x})$ 的以下结果颇有用处.

4.1.4 定理　设 $K = Y^\circ + \mathbf{R}g(\bar{x})$，则有以下结论：

(i) $F_G(\bar{x}) \supset A^{-1}K$.

(ii) 若 $A^{-1}K \neq \emptyset$，则 $F_G(\bar{x}) = A^{-1}K, F_G^*(\bar{x}) = -A^*Q$.

证　(i) 若 $z \in A^{-1}K$，则 $\exists r \in \mathbf{R}: Az \ll rg(\bar{x})$. 于是当 $0 < t \downarrow 0, w \to z$ 时，

$$g(\bar{x} + tw) = g(\bar{x}) + tAz + o(t) \leqslant (1 + rt)g(\bar{x}) \leqslant 0,$$

即 $\bar{x} + tw \in G$. 可见 $z \in F_G(\bar{x})$.

(ii) 设 $z \in F_G(\bar{x})$. 若 $Az \overline{\in} K$，则由 1.1.2 有 $0 \neq \lambda \in Y^*: \langle \lambda, Az \rangle \geqslant \langle \lambda, K \rangle$，这推出 $\langle \lambda, Az \rangle \geqslant 0, \lambda \in -K^* = Q$. 取模充分小的 $x \in A^{-1}K$；取 $y \in Y_-^\circ, r \in \mathbf{R}$，使 $Ax = y + rg(\bar{x})$. 因 $z \in F_G(\bar{x})$，当 $0 < t \downarrow 0$ 时有 $g(\bar{x} + t(z - x)) \leqslant 0$，于是

$$\begin{aligned} 0 &\leqslant \langle \lambda, tAz - g(\bar{x} + t(z - x)) \rangle \\ &= \langle \lambda, tAx - g(\bar{x}) + o(t) \rangle \\ &= \langle \lambda, ty + o(t) \rangle < 0, \end{aligned}$$

得出矛盾. 因此 $z \in A^{-1}K$. 故得 $F_G(\bar{x}) = A^{-1}K$. 进而用 3.1.2 得 $F_G^*(\bar{x}) = A^*K^* = -A^*Q$. □

4.1.5 定理　设 f 在 $\bar{x}$ 可微，$T = f'(\bar{x}) \neq 0, G = \{x : f(x) \leqslant f(\bar{x})\}, H = \{x : f(x) = f(\bar{x})\}, K = \{z : Tz < 0\}$. 则 $\overline{K} = \{z : Tz \leqslant 0\}$；

$$V_G(\bar{x}) \subset F_G(\bar{x}) = K, \ V_G^*(x) \supset F_G^*(\bar{x}) = \mathbf{R}_- T; \qquad (1)$$

$$T_G(\bar{x}) \subset K_G(\bar{x}) \subset \overline{K}, \ T_G^*(\bar{x}) \supset K_G^*(\bar{x}) \supset \mathbf{R}_- T; \qquad (2)$$

$$K_H(\bar{x}) \subset N(T), K_H^*(\bar{x}) \supset \mathbf{R}T. \qquad (3)$$

若 $f \in C^1$，则 (1) ~ (3) 中的包含皆为等式.

证　直接看出 $\overline{K} = \{z : Tz \leqslant 0\}, (\overline{K})^* = K^* = \mathbf{R}_- T$. 分别从 4.1.4，4.1.2 与 4.1.3 推出 (1)，(2) 与 (3). 下面设 $f \in C^1$. 若 $z \in K$，则当 $t \downarrow 0, x \xrightarrow{G} \bar{x}, w \to z$ 时，有

$$f(x + tw) = f(x) + \langle f'(x + \tau w), tw \rangle$$

$$\leqslant f(\bar{x})\ (0 < \tau < t),$$

即 $x + tw \in G$. 可见 $z \in V_G(\bar{x})$. 因此(1) 中的包含是等式. 其次，由 $\overline{K} = \overline{V_G(\bar{x})} \subset T_G(\bar{x})$ 知(2) 中的包含是等式. 直接由 4.1.3 知(3) 是等式. □

现在转向非光滑的情况，即仅假定 f,g,h 在 $\bar{x}$ 邻近为 Lip. 在这种情况下，要得出 4.1.1 ~ 4.1.5 的完全的推广是困难的. 下面在一定附加条件下得出 4.1.1 ~ 4.1.5 的某种类似. 为记号简便，约定 $g_\lambda = \lambda g, h_\mu = \mu h (\lambda \in Y^*, \mu \in Z^*)$. 令

$$K = \{z : g_\lambda^\circ(\bar{x},z) \leqslant 0, h_\mu^\circ(\bar{x},z) = 0 (\forall \lambda \in Q, \mu \in Z^*)\}; \quad (4)$$

$$\Phi = \bigcup \{\partial g_\lambda(\bar{x}) + \partial h_\mu(\bar{x}) : \lambda \in Q, \mu \in Z^*\}. \quad (5)$$

K 与 Φ 分别对应光滑情形下的集 $(A^{-1}K_x) \cap N(B)$ 与 $A^*Q + R(B^*)$.

4.1.6 定理 若 $\forall (\lambda,\mu) \in Q \times Z^*$: g_λ 与 h_μ 在 $\bar{x}$ 正则，则 $K_M(\bar{x}) \subset K$, $-K_M^*(\bar{x}) \supset \Phi$ (与 4.1.1 对照).

证 任给 $z \in K_M(\bar{x})$. 取 $t_n \downarrow 0, z_n \to z$, 使 $x_n \overset{\text{def}}{=} \bar{x} + t_n z_n \in M$. 任给 $\lambda \in Q$, 有

$$g_\lambda^\circ(\bar{x},z) = D_+ g_\lambda(\bar{x},z) = \lim_n \frac{g_\lambda(\bar{x} + t_n z) - g_\lambda(\bar{x})}{t_n}$$

$$\leqslant \limsup_n \frac{g_\lambda(\bar{x} + t_n z) - g_\lambda(x_n)}{t_n} = 0.$$

同理有 $h_\mu^\circ(\bar{x},z) = 0 (\forall \mu \in Z^*)$, 因此 $z \in K$. 其次，当 $z \in K$, $(\lambda,\mu) \in Q \times Z^*$, $u \in \partial g_\lambda(\bar{x})$, $v \in \partial h_\mu(\bar{x})$ 时，有

$$u(z) + v(z) \leqslant g_\lambda^\circ(\bar{x},z) + h_\mu^\circ(\bar{x},z) \leqslant 0.$$

可见 $\langle \Phi, K \rangle \leqslant 0$, 因此 $\Phi \subset -K^* \subset -K_M^*(x)$. □

4.1.7 推论 若 $\forall \lambda \in Q$: g_λ 在 $\bar{x}$ 正则，则

$$K_G(\bar{x}) \subset \{z : g_\lambda^\circ(\bar{x},z) \leqslant 0 (\forall \lambda \in Q)\},$$

$$-K_G^*(\bar{x}) \supset \bigcup_{\lambda \in Q} \partial g_\lambda(\bar{x}).$$

若 $\forall \mu \in Z^*$: h_μ 在 $\bar{x}$ 正则，则

$$K_H(\bar{x}) \supset \{z : h_\mu^\circ(\bar{x},z) = 0(\forall \mu \in Z^*)\},$$

$$-K_H^*(x) \supset \bigcup_{\mu \in Z^*} \partial h_\mu(\bar{x}).$$

4.1.8 定理 设 $Y = \mathbf{R}^m$；$\forall u_i \in \partial g_i(\bar{x})$，$\{u_i : i \in I\}$ 线性无关；$K = \{z : g_i^\circ(\bar{x},z) < 0(\forall i \in I)\}$，$\Phi = \sum_{i\in I} \mathbf{R}_- \partial g_i(\bar{x})$. 则 $\overline{K} = \{z : g_i^\circ(\bar{x},z) \leqslant 0(\forall i \in I)\}$；$K^* = (\overline{K})^* = \Phi$；

$$K \subset V_G(\bar{x}) \subset F_G(\bar{x}),\ \Phi \supset V_G^*(\bar{x}) \supset F_G^*(\bar{x}); \qquad (6)$$

$$\overline{K} \subset T_G(\bar{x}) \subset K_G(\bar{x}),\ \Phi \supset T_G^*(\bar{x}) \supset K_G^*(\bar{x}). \qquad (7)$$

若 $g_i(i \in I)$ 皆在 $\bar{x}$ 正则，则式(6)，式(7)中的包含为等式.

证 不妨设 $g(\bar{x}) = 0$，因而 $I = \{1,2,\cdots,m\}$.

f 必 $K \neq \varnothing$. 否则，问题

$$\exists z \in X ; g_i^\circ(\bar{x},z) < 0 \quad (1 \leqslant i \leqslant m)$$

无解. 于是由 3.3.2 有 $0 \neq \lambda \in \mathbf{R}_+^m : \sum \lambda_i g_i^\circ(\bar{x},z) \geqslant 0(\forall z \in X)$，这推出 $0 \in \sum \lambda_i \partial g_i(\bar{x})$(2.2.1，2.1.5，2.1.6)，与定理中的线性无关假定矛盾.

令 $D = \{z : g_i^\circ(\bar{x},z) \leqslant 0(1 \leqslant i \leqslant m)\}$，则 D 闭且 $K \subset D$，因此 $\overline{K} \subset D$. 任给 $z \in D$，取 $z_0 \in K$，令 $z_t = t'z_0 + tz(0 < t < 1)$. 由 $g_i^\circ(x,\cdot)$ 凸得

$$g_i^\circ(\bar{x},z_t) \leqslant t' g_i^\circ(\bar{x},z_0) + t g_i^\circ(\bar{x},z) < 0 \quad (1 \leqslant i \leqslant m),$$

可见 $z_t \in K$. 于是 $z_t \to z \in \overline{K}(t \uparrow 1)$，因此 $\overline{K} = D$.

易见 $\Phi \subset K^* = (\overline{K})^*$. 今证 $\Phi \supset K^*$，首先证 Φ 弱*闭. 设 $\lambda_\alpha u_\alpha \xrightarrow{*} v$，$\lambda_\alpha u_\alpha = \sum_i \lambda_{\alpha i} u_{\alpha i} \in \Phi$，$\lambda_{\alpha i} \in \mathbf{R}_-$，$u_{\alpha i} \in \partial g_i(x)$. 因 $\partial g_i(\bar{x})$ 弱*紧，不妨设 $u_{\alpha i} \xrightarrow{*} u_i \in \partial g_i(\bar{x})$. 令 $u = (u_i)$. 若 $\{\lambda_\alpha\}$ 无界，可设 $|\lambda_\alpha| \to \infty$. 令 $\lambda'_\alpha = \lambda_\alpha / |\lambda_\alpha|$，不妨设 $\lambda'_\alpha \to \lambda = (\lambda_i) \in \mathbf{R}_-^m$，则 $\lambda u = \sum \lambda_i u_i = 0$，这与 $\{u_i\}$ 线性无关矛盾. 故 $\{\lambda_\alpha\}$ 有界，因而不妨设 $\lambda_\alpha \to \lambda \in \mathbf{R}_-^m$. 于是 $v = \sum \lambda_i u_i \in \Phi$，$\Phi$ 为弱*闭. 若 $w \in X^* \backslash \Phi$，则由 1.1.3 有 $z \in X ; w(z) < \langle \Phi,z \rangle$. 这推出 $w(z) < 0$，$\langle \Phi,z \rangle \geqslant 0$；后者推出

$z \in \overline{K}$，因此 $w \overline{\in} K^*$. 这就证得 $\Phi \supset K^*$，从而 $K^* = (\overline{K})^* = \Phi$.

2 若 $z \in K$，则当 $0 < t \downarrow 0, x \xrightarrow{G} \overline{x}, w \to z$ 时，有

$$
\begin{aligned}
g_i(x + tw) &\leqslant g_i(x + tw) - g_i(x) \\
&= \Delta g_i(x, tz) + g_i(x + tw) - g_i(x + tz) \\
&\leqslant t g_i^\circ(\overline{x}, z) + o(t) \leqslant 0 \quad (1 \leqslant i \leqslant m),
\end{aligned}
$$

故 $x + tw \in G$，从而 $z \in V_G(\overline{x})$. 这证得式(6). 进而有 $\overline{K} \subset \overline{V_G(\overline{x})} \subset T_G(\overline{x})$，于是(7) 成立.

3 设 $g_i (1 \leqslant i \leqslant m)$ 在 $\overline{x}$ 正则. 任给 $z \in K_G(\overline{x})$，取 $t_n \downarrow 0$，$z_n \to z$，使 $x_n \stackrel{\text{def}}{=} \overline{x} + t_n z_n \in G$，则

$$
\begin{aligned}
g_i^\circ(\overline{x}, z) = \mathrm{D}_+ g_i(\overline{x}, z) &= \lim_n \frac{g_i(\overline{x} + t_n z) - g_i(\overline{x})}{t_n} \\
&\leqslant \limsup_n \frac{g_i(\overline{x} + t_n z) - g_i(x_n)}{t_n} = 0 \ (1 \leqslant i \leqslant m),
\end{aligned}
$$

可见 $z \in \overline{K}$. 因此式(7) 为等式.

任给 $z \in F_G(\overline{x})$，有 $z \in K_G(\overline{x}) = \overline{K}$. 若 $z \overline{\in} K$，则对某个 i 有 $g_i^\circ(\overline{x}, z) = 0$. 取 $u \in \partial g_i(\overline{x})$，使 $u(z) = 0$. 由定理的线性无关假设，必定 $u \neq 0$；于是有 $z_n \to z$，使 $u(z_n) > 0$. 另一方面，由 $z \in F_G(\overline{x})$，可设当 $t \downarrow 0$ 时 $g(\overline{x} + t z_n) \leqslant 0$. 于是

$$
u(z_n) \leqslant g_i^\circ(\overline{x}, z_n) = \lim_{t \downarrow 0} \frac{g_i(\overline{x} + t z_n) - g_i(\overline{x})}{t} \leqslant 0,
$$

得出矛盾. 因此 $z \in K$. 这证得式(6) 为等式. □

以 f 代 g，从 4.1.8 得出下面推论.

4.1.9 推论 设 $0 \overline{\in} \partial f(\overline{x})$；$G = \{x : f(x) \leqslant f(\overline{x})\}$，$K = \{z : f^\circ(\overline{x}, z) < 0\}$. 则

$$
\overline{K} = \{z : f^\circ(\overline{x}, z) \leqslant 0\}, K^* = (\overline{K})^* = \mathbf{R}_+ \partial f(\overline{x});
$$

$$
K \subset V_G(\overline{x}) \subset F_G(\overline{x}), \ \mathbf{R}_+ \partial f(\overline{x}) \supset V_G^*(\overline{x}) \supset F_G^*(\overline{x}); \quad (8)
$$

$$
\overline{K} \subset T_G(\overline{x}) \subset K_G(\overline{x}), \ \mathbf{R}_+ \partial f(\overline{x}) \supset T_G^*(\overline{x}) \supset K_G^*(\overline{x}). \quad (9)
$$

若 f 在 $\overline{x}$ 正则，则(8)，(9)中的包含为等式.

4.1.9 可与 4.1.5 相对照.

§4.2 Fritz John 定理

本节考虑问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D. \tag{1}$$

设 $\bar{x} \in M = D \cap G \cap H$，$f, g, h$ 在 $\bar{x}$ 处可微，$K_g = Y_- + \mathbf{R}g(\bar{x})$，$Q = -K_g^* - Y_+^* \cap \{g(\bar{x})\}^\perp$，$L(\cdot, \rho, \lambda, \mu) = \rho f + \lambda g + \mu h$. 注意 $L_x(\bar{x}, \rho, \lambda, \mu) = \rho\Gamma + A^*\lambda + B^*\mu$. 令

$$F_f(\bar{x}) = \{z: \text{当 } t \downarrow 0, w \to z \text{ 时 } f(\bar{x} + tw) < f(\bar{x})\}, \tag{2}$$

称它为 f 在 $\bar{x}$ 的**降锥**. $F_f(\bar{x})$ 显然是开锥，且当 $f'(\bar{x}) \neq 0$ 时 $F_f(\bar{x}) = \{z: \langle f'(\bar{x}), z\rangle < 0\}$.

4.2.1 引理　若 $\bar{x}$ 是(1)的局部最优解，则

$$F_f(\bar{x}) \cap K_M(\bar{x}) = \varnothing.$$

证　任给 $z \in K_M(\bar{x})$，$\exists t_n \downarrow 0, z_n \to z: x_n \stackrel{\text{def}}{=} \bar{x} + t_n z_n \in M$，于是 $f(\bar{x} + t_n z_n) \geqslant f(\bar{x})$，这表明 $z \overline{\in} F_f(\bar{x})$. □

为行文简便起见，约定以下术语：若 $\{K_j\}$ 是一组锥，存在不全为零的 $u_j \in K_j^*$，使 $\sum u_j = 0$，则说 $\{K_j^*\}$ 为 SOS(依[223]).

4.2.2 Dubovickii-Miljutin 引理　设 $K_j (0 \leqslant j \leqslant n)$ 是非空凸锥，当 $j < n$ 时 K_j 是开的，则 $\bigcap K_j = \varnothing \Leftrightarrow \{K_j^*\}$ 为 SOS.

证　若 $\bigcap K_j = \varnothing$，则 $\exists m < n: K \stackrel{\text{def}}{=} \bigcap_{j \leqslant m} K_j \neq \varnothing$，$K \cap K_{m+1} = \varnothing$. 因 K 是非空开凸锥，故由 1.1.2 有 $u \in X^*: u(K) > u(K_{m+1})$，这得出 $u \in K^* = \sum_{j \leqslant m} K_j^*$ (1.4.7)，$-u \in K_{m+1}^*$. 由此易见 $\{K_j^*\}$ 为 SOS. 反之，若 $\{K_j^*\}$ 为 SOS，则有 $u_j \in K_j^* (0 \leqslant j \leqslant n)$，使 $\sum u_j = 0$，且某个 $u_i \neq 0, i < n$. 于是必 $\bigcap K_j = \varnothing$，否则有 $x \in \bigcap K_j$，但这推出 $0 = \sum u_j(x) \geqslant u_i(x) > 0$ (用 1.4.2(ii))！ □

以下是本节的中心结果.

4.2.3 Fritz John 定理　设 $Y^{\circ}_{+} \neq \varnothing$, $K \subset F_D(\bar{x})$ 是一非空开凸锥；h 在 $\bar{x}$ 邻近为 C^1 映射，$R(B)$ 为闭集. 若 $\bar{x}$ 是问题(1) 的局部最优解，则有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+ \times Q \times Z^*$，使

$$L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) \in K^*.$$

证　不妨设 $f'(\bar{x}) \neq 0$（否则可取 $\bar{\rho} = 1, \bar{\lambda} = 0, \bar{\mu} = 0$），$K_1 \stackrel{\mathrm{def}}{=} A^{-1}(Y^{\circ}_{+} + \mathbf{R}g(\bar{x})) \neq \varnothing$（否则 $0 \,\overline{\in}\, R(A) + Y^{\circ}_{+} + \mathbf{R}g(\bar{x})$，于是由 1.1.2 有 $0 \neq \bar{\lambda} \in Y^*$：$0 \leqslant \langle \bar{\lambda}, R(A) + Y^*_+ + \mathbf{R}g(\bar{x}) \rangle$，因而 $\bar{\lambda} \in Q$，$A^* \bar{\lambda} = 0$，可取 $\bar{\rho} = 0, \bar{\mu} = 0$），$R(B) = Z$（否则有 $0 \neq \bar{\mu} \in N(B^*)$，于是可取 $\bar{\rho} = 0, \bar{\lambda} = 0$）. 令 $K_0 = F_f(\bar{x})$，$K_2 = K$，$K_3 = K_H(\bar{x}) = N(B)$(4.1.3). 则 $K_j (0 \leqslant j \leqslant 3)$ 为非空凸锥，K_0, K_1, K_2 是开的；$K_0^* = \mathbf{R}_- f'(\bar{x})$；$K_1^* = A^*(Y^{\circ}_{-} + \mathbf{R}g(\bar{x}))^* = -A^*Q$ (3.1.2)；$K_3^* = R(B^*)$(4.1.3). 由 $K_1 \subset F_G(\bar{x})$(4.1.4)，$K_2 \subset F_D(\bar{x})$ 及 1.4.7，4.2.1 有

$$\bigcap_{j=0}^{3} K_j \subset F_f(\bar{x}) \cap F_G(\bar{x}) \cap F_D(\bar{x}) \cap K_H(\bar{x})$$

$$\subset F_f(\bar{x}) \cap K_M(\bar{x}) = \varnothing.$$

于是由 4.2.2 有不全为零的 $u_j \in K_j^* (0 \leqslant j \leqslant 3)$：$\sum u_j = 0$. 设 $u_0 = -\rho f'(x)$，$u_1 = -A^* \lambda$，$u_3 = -B^* \bar{\mu}$，$(\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+ \times Q \times Z^*$，则必定 $(\bar{\rho}, \bar{\lambda}, \bar{\mu}) \neq 0$，

$$L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) = -u_0 - u_1 - u_3 = u_2 \in K_2^* = K^*. \qquad \square$$

注　若 D 凸且 $D \neq \varnothing$，则 $F_D(\bar{x}) = \mathrm{cone}(D - \bar{x})$ 是非空开凸锥，在 4.2.3 中取 $K = F_D(\bar{x})$ 得

$$L_x(\bar{x}, \bar{\rho}, \lambda, \bar{\mu}) \in F_D^*(\bar{x}) = (D - \bar{x})^*.$$

若 $\bar{x} \in D$，则必 $F_D(\bar{x}) = X$，从而 $F_D^*(\bar{x}) = \{0\}$，于是

$$L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) = 0.$$

若 $Y = \mathbf{R}^m$，$I = \{i : g_i(\bar{x}) = 0\}$，则易验证 $Q = \mathbf{R}^I_+$，于是由 4.2.3 得到下面的推论.

4.2.4 推论　若在 4.2.3 的条件下，再假定 $Y = \mathbf{R}^m$，则有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+ \times \mathbf{R}^I_+ \times Z^*$ 使 $L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) \in K^*$.

通常所说的 Fritz John 定理，乃指 4.2.4 而言.

今将 4.2.3 用到问题(1)的特款：

$$\min f(x),\ g(x) \leqslant 0, x \in D \tag{3}$$

与

$$\min f(x),\ h(x) = 0, x \in D. \tag{4}$$

首先，取 $Y = \{0\}$，直接从 4.2.3 得出推论.

4.2.5 推论 设 h 在 $\bar{x}$ 邻近为 C^1 映射，$R(B)$ 为闭集；$K \subset F_D(\bar{x})$ 是一非空开凸锥. 若 $\bar{x}$ 是问题(4) 的局部最优解，则有 $0 \neq (\bar{\rho}, \bar{\mu}) \in \mathbf{R}_+ \times Z^*$，使 $\bar{\rho} f'(\bar{x}) + B^* \bar{\mu} \in K^*$.

对于问题(3) 有如下较强的结论.

4.2.6 定理 设 $K \subset F_D(\bar{x})$ 是一非空开凸锥，$\bar{x}$ 是问题(3) 的局部最优解. 若(i) $Y_+^\circ \neq \varnothing$，或(ii) $g \in C^1$，$R(A) + K_g = Y$，则有 $0 \neq (\bar{\rho}, \bar{\lambda}) \in \mathbf{R}_+ \times Q$，使 $\bar{\rho} f'(\bar{x}) + A^* \bar{\lambda} \in K^*$.

证 当 $Y_+^\circ \neq \varnothing$ 时结论直接由 4.2.3 推出. 下面设条件(ii) 满足. 如同证 4.2.3 一样，可设 $f'(\bar{x}) \neq 0$. 令 $K_0 = F_f(\bar{x})$，$K_1 = K$，$K_2 = A^{-1} K_g$，则 K_0, K_1 是非空开凸锥；条件(ii) 推出 K_2 为非空凸锥，$K_2 \subset K_G(\bar{x})$，$K_2^* = -A^* Q$(4.1.2). 因

$$\begin{aligned} K_0 \cap K_1 \cap K_2 &\subset F_f(\bar{x}) \cap F_D(\bar{x}) \cap K_G(\bar{x}) \\ &\subset F_f(\bar{x}) \cap K_M(\bar{x}) = \varnothing, \end{aligned}$$

故由 4.2.2 有 $0 \neq (u_0, u_1, u_2) \in K_0^* \times K_1^* \times K_2^*$；$u_0 + u_1 + u_2 = 0$. 设 $u_0 = -\bar{\rho} f'(\bar{x})$，$\bar{\rho} \in \mathbf{R}_+$，$u_2 = -A^* \bar{\lambda}$，$\bar{\lambda} \in Q$，则 $(\bar{\rho}, \bar{\lambda}) \neq 0$，$\bar{\rho} f'(\bar{x}) + A^* \bar{\lambda} = -u_0 - u_2 = u_1 \in K_1^* = K^*$. □

关于 4.2.6 的条件(ii) 可参考[128,149]. 用条件(ii) 替代 $Y_+^\circ \neq \varnothing$，在无限维空间中是极有意义的. [128]对无限维最优化问题提供了一系列最优性条件.

问题(1)可抽象成一个更一般的约束极小问题：

$$\min f(x),\ x \in M_j, \quad j = 1, 2, \cdots, n. \tag{5}$$

若取 $D = M_1$，$G = M_2$，$H = M_3$，则从问题(5)得出(1). 关于问题(5)有以下著名结果.

4.2.7 Dubovickii-Miljutin 定理(1965)　设 $M_j \subset X, \bar{x} \in M \overset{\text{def}}{=} \bigcap M_j, K_0 = F_f(\bar{x}), K_j = F(M_j, \bar{x})(1 \leqslant j < n), K_n = K(M_n, \bar{x})$,假定 $K_j(0 \leqslant j \leqslant n)$ 非空凸.

(i) 若 $\bar{x}$ 是问题(5) 的局部最优解,则 $\{K_j^* : 0 \leqslant j \leqslant n\}$ 为 SOS.

(ii) 若 $\{K_j^* : 0 \leqslant j \leqslant n\}$ 为 SOS,M_j 凸,

$$N \overset{\text{def}}{=} M_1 \cap \cdots \cap M_{n-1} \cap M_n \neq \varnothing,$$

f 严格拟凸且为 usc,则 $\bar{x}$ 是问题(5) 的最优解.

证　(i) 直接从 4.2.1 与 4.2.2 推出.

(ii) 若 $\bar{x}$ 非最优解,则 $\exists x_0 \in M: f(x_0) < f(\bar{x})$. 取 $x \in N$,令 $x_\tau = \tau' x_0 + \tau x, \tau > 0$ 充分小;$z_\tau - x_\tau - \bar{x}$,则 $\forall t \in (0,1): \bar{x} + tz_\tau = t'\bar{x} + tx_\tau \in N$,这推出 $z_\tau \in \bigcap_1^n K_j$. 因 f 为 usc,故当 $z \to z_\tau$ 时(这意味着 $\bar{x} + z$ 邻近 x_0),$f(\bar{x} + z) < f(\bar{x})$. 由 f 严格拟凸得出

$$f(\bar{x} + tz) < f(\bar{x})(0 < t < 1, z \to z_\tau),$$

因此 $z_\tau \in K_0$,从而 $\bigcap_0^n K_j \neq \varnothing$,这与 $\{K_j^*\}$ 为 SOS 矛盾. □

4.2.7 是 60 年代寻求统一的最优性条件的种种努力中最典型的结果,它可看作经典的 Lagrange 乘子定理的抽象形式:$\{K_j^*\}$ 为 SOS 意味着有 $0 \neq (u_j) \in \coprod K_j^*$ 使 $\sum u_j = 0$,"$\sum u_j = 0$" 可看作抽象的"Euler-Lagrange 方程",而 u_j 似乎是抽象的 **Lagrange 乘子**. 不过,4.2.7 的有效应用有赖于方程 $\sum u_j = 0$ 的具体化. 4.2.3～4.2.6 正提供了某种具体化.

参考文献　[11,12,61,93,125,128,131,132,149,203,223,238].

§4.3　Kuhn-Tucker 条件

本节考虑最优化问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0. \tag{1}$$

设 $\bar{x} \in M = G \cap H, L(\cdot, \lambda, \mu) = f + \lambda g + \mu h$. 假定 f, g, h 在 $\bar{x}$ 可微.

4.3.1 定义 若存在 $(\bar{\lambda}, \bar{\mu}) \in Q \times Z^*$，使得 $L_x(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$，则说问题(1) 在点 $\bar{x}$ 满足 Kuhn-Tucker 条件，或称 $\bar{x}$ 为(1) 的 Kuhn-Tucker 点，称 $\bar{\lambda}, \bar{\mu}$ 为 Kuhn-Tucker 乘子(分别简称为 **K-T 条件**、**K-T 点**与 **K-T 乘子**).

形式上，4.3.1 并未涉及 $\bar{x}$ 是否为最优解. 从 Kuhn，Tucker 开始的一系列研究致力于加于约束的一定条件 —— 通常称为约束品性 —— 用以保证局部最优点是 K-T 点.

首先注意，若依 4.2.3 得出 $L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) = 0$ 且 $\bar{\rho} \neq 0$，则 $\bar{\lambda}/\bar{\rho}$，$\bar{\mu}/\bar{\rho}$ 就是 K-T 乘子. 据此可证：

4.3.2 定理 设 $Y_+^{*\circ} \neq \varnothing$，$h$ 在 $\bar{x}$ 邻近为 C^1 映射，$R(B) = Z$. 若 $\bar{x}$ 是问题(1) 的局部最优点，则以下每个条件蕴涵 $\bar{x}$ 为 K-T 点：

(i) $S \overset{\text{def}}{=\!=} (A, B) : X \to Y \times Z$ 是满射；

(ii) $N(B) \cap A^{-1}(Y_-^\circ + \mathbf{R}g(\bar{x})) \neq \varnothing$；

(iii) g 为 * 伪凸(即 $\forall \lambda \in Y_+^*$；$\lambda g$ 为伪凸)，$h(x) = Bx - b$，且 $\exists \hat{x}$：$g(\hat{x}) \ll 0, h(\hat{x}) = 0$.

证 由 4.2.3，有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+ \times Q \times Z^*$：$L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) = 0$. 设 $\rho = 0$，则 $0 \neq (\bar{\lambda}, \bar{\mu}) \in N(S^*)$. 必定 $\bar{\lambda} \neq 0$(否则 $0 \neq \bar{\mu} \in N(B^*)$，与 $R(B) = Z$ 矛盾). 若条件(i) 满足，则 $N(S^*) = R(S)^\perp = \{0\}$，这与 $(\bar{\lambda}, \bar{\mu}) \neq 0$ 矛盾. 若条件(ii) 满足，则 $\exists z \in N(B)$：$Az \in Y_-^\circ + \mathbf{R}g(\bar{x})$，于是

$$0 = \langle A^*\bar{\lambda} + B^*\bar{\mu}, z\rangle = \langle \bar{\lambda}, Az\rangle < 0,$$

得出矛盾. 若条件(iii)满足，令 $z = \hat{x} - \bar{x}$，则 $Bz = 0$，

$$\langle \bar{\lambda}g'(\bar{x}), \hat{x} - \bar{x}\rangle = \langle A^*\bar{\lambda}, z\rangle = \langle A^*\bar{\lambda} + B^*\bar{\mu}, z\rangle = 0,$$

于是由 $\bar{\lambda}g$ 伪凸推出 $0 = \langle \bar{\lambda}, g(\bar{x})\rangle \leqslant \langle \bar{\lambda}, g(\hat{x})\rangle < 0$，亦得矛盾. 故必 $\bar{\rho} \neq 0$，从而 $\bar{x}$ 为 K-T 点. □

4.3.3 推论 设 $Y = \mathbf{R}^m$，h 在 $\bar{x}$ 邻近为 C^1 映射，$R(B) = Z$.

若 $\bar{x}$ 是问题(1) 的局部最优解，则以下每个条件蕴涵 $\bar{x}$ 为 K-T 点：

(i) $(A_I, B): X \to \mathbf{R}^I \times Z$ 为满射；

(ii) $N(B) \cap A_I^{-1}\mathbf{R}_-^{\circ I} \neq \varnothing$；

(iii) g_I 为 * 伪凸(或 g_i 伪凸，$i \in I$)，$h(x) = Bx - b$，$\exists \hat{x}: g_I(\hat{x}) \ll 0, h(\hat{x}) = 0$.

证 以 g_I 代 g，从 4.3.2 推出. □

注 若 $Y = \mathbf{R}^m$，则 $\bar{x}$ 为 K-T 点意味着 $\exists \bar{\lambda}_i \geqslant 0 (i \in I), \bar{\mu} \in Z^*$，使得

$$f'(\bar{x}) + \sum_{i \in I} \bar{\lambda}_i g'_i(\bar{x}) + B^* \bar{\mu} = 0.$$

若再设 $Z = \mathbf{R}^p$，则 4.3.3 中条件(i) 等价于 $\{g'_i(\bar{x}), h'_j(\bar{x}) : i \in I, 1 \leqslant j \leqslant p\}$ 线性无关.

得出 K-T 点的另一途径基于以下事实：$\bar{x}$ 是问题(1) 的 K-T 点 $\Leftrightarrow \exists (\bar{\lambda}, \bar{\mu}) \in Q \times Z^*: -f'(\bar{x}) = A^* \bar{\lambda} + B^* \bar{\mu} \Leftrightarrow -f'(\bar{x}) \in A^*Q + R(B^*)$. 因此，若有 $K \subset X$ 满足

$$K \cap F_f(\bar{x}) = \varnothing, \ -K^* \subset A^*Q + R(B^*), \qquad (2)$$

则 $\bar{x}$ 必为 K-T 点. 否则 $f'(\bar{x}) \bar{\in} K^*$，于是 $\exists z \in K: \langle f'(\bar{x}), z \rangle < 0$，从而 $z \in K \cap F_f(\bar{x})$，与 $K \cap F_f(\bar{x}) = \varnothing$ 矛盾. 若 $\bar{x}$ 是(1) 的局部最优解，则 $K_M(\bar{x}) \cap F_f(\bar{x}) = \varnothing$(4.2.1)；其次，在一定条件下有 $-K_M^*(\bar{x}) = A^*Q + R(B^*)$(4.1.1). 因此取 $K = K_M(\bar{x})$ 是一种可考虑的选择. 从以上分析引出：

4.3.4 定理 设 $\bar{x}$ 是问题(1) 的局部最优解，则以下每个条件蕴涵 $\bar{x}$ 为 K-T 点：

(i) $-K_M(\bar{x}) = N(B) \cap A^{-1}(Y \cap Q^*)$，$A^*Q + R(B^*)$) 弱* 闭；

(ii) $g, h \in C^1$，$(A, B): X \to Y \times Z$ 是满射，K_g 闭且 $A^*Q + R(B^*)$ 弱* 闭；

(iii) $K_M(\bar{x}) = N(B) \cap A^{-1}K_g$，$K_g$ 与 $R(B)$ 闭，$R(A) + K_g = Y$ 且 $N(B) + A^{-1}K_g = X$.

证 (i) 4.1.1 (ii) 之证实际上已得出:当 $A^*Q+R(B^*)$ 弱*闭时 $(N(B)\cap A^{-1}(Y\cap Q^*))^*\subset A^*Q+R(B^*)$,而这得出

$$-K_M^*(\bar{x})=A^*Q+R(B^*).$$

(ii) 直接由 4.1.1(ii)得出.

(iii) 在所给条件下相继应用 3.1.8,1.1.8 与 3.1.2 得

$$\begin{aligned}-K_M^*(\bar{x})&=(N(B)\cap A^{-1}(-K_g))^*\\&=N(B)^{\perp}-(A^{-1}K_g)^*\\&=R(B^*)-A^*K_g^*\\&=A^*Q+R(B^*).\end{aligned}$$ □

注 与 4.3.2 比较,4.3.4 的优点是其中未假定 $Y_+^\circ\neq\varnothing$. 4.3.4 的缺点是其中出现不易确定的 $K_M(\bar{x})$. 不过,4.3.4 中的条件

$$-K_M(\bar{x})=N(B)\cap A^{-1}(Y\cap Q^*)$$

与
$$K_M(\bar{x})=N(B)\cap A^{-1}K_g$$

可分别改成

$$N(B)\cap A^{-1}(Y\cap Q_-^*)\cap F_f(\bar{x})=\varnothing$$

与
$$N(B)\cap A^{-1}K_g\cap F_f(\bar{x})=\varnothing.$$

4.3.5 推论 设 $Y=\mathbf{R}^m$,$\bar{x}$ 是问题(1) 的局部最优解,则以下每个条件蕴涵 $\bar{x}$ 为 K-T 点:

(i) $N(B)\cap A^{-1}\mathbf{R}_-^I\cap F_f(\bar{x})=\varnothing$,$A_I^*\mathbf{R}_+^I+R(B^*)$ 弱* 闭(当 $\dim X<\infty$ 时后者不需验证);

(ii) $N(B)\cap A^{-1}\mathbf{R}_-^I\cap F_f(\bar{x})=\varnothing$,$R(B)$ 闭,$R(A_I)+\mathbf{R}_+^I=\mathbf{R}^I$,$N(B)+A_I^{-1}\mathbf{R}_+^I=X$.

现在将以上结果用到问题(1)的特款

$$\min f(x),\ g(x)\leqslant 0 \tag{3}$$

与
$$\min f(x),\ h(x)=0. \tag{4}$$

注意可行点 $\bar{x}$ 是(3) 与(4) 的 K-T 点分别意味着 $f'(\bar{x})\in -A^*Q$ 与 $f'(\bar{x})\in R(B^*)$. 由 4.3.2 ~ 4.3.5 推出:

4.3.6 推论 设 $\bar{x}$ 是问题(3)的局部最优解，则以下每个条件蕴涵 $\bar{x}$ 为 K-T 点：

(i) $Y_-^\circ \neq \varnothing, R(A) = Y$;

(ii) $A^{-1}(Y_-^\circ + \mathbf{R}g(\bar{x})) \neq \varnothing$;

(iii) g 为 * 伪凸，$\exists \hat{x}: g(\hat{x}) \ll 0$;

(iv) $A^{-1}(Y \cap Q_-^*) \cap F_f(\bar{x}) = \varnothing, A^*Q$ 弱*闭；

(v) $g \in C^1, R(A) + K_g = Y$;

(vi) $Y = \mathbf{R}^m, R(A_I) = \mathbf{R}^I$;

(vii) $Y = \mathbf{R}^m, A_I^{-1}\mathbf{R}_+^{\circ I} \neq \varnothing$;

(*viii*) $Y = \mathbf{R}^m, A_I^{-1}\mathbf{R}_-^I \cap F_f(\bar{x}) = \varnothing$.

需要说明的只有(v). 当条件(v)满足时，由 4.1.2(ii) 有 $A^{-1}K_g \subset K_G(\bar{x}), (A^{-1}K_g)^* = -A^*Q$. 于是 $K \overset{\text{def}}{=} A^{-1}K_g$ 满足条件(2)(取 $B = 0$).

4.3.7 推论 设 $\bar{x}$ 是问题(4)的局部最优解，则以下每个条件蕴涵 $\bar{x}$ 为 K-T 点：

(i) $h \in C^1, R(B) = Z$;

(ii) $N(B) \cap F_f(\bar{x}) = \varnothing, R(B)$ 为闭集.

以上结论的直接证明亦很简单. 例如，若条件(ii)满足，而 $f'(\bar{x}) \notin R(B^*)$，则因 $R(B^*) = N(B)^{\perp}$ (1.1.8)，必有 $z \in N(B)$: $\langle f'(\bar{x}), z \rangle < 0$，于是 $z \in N(B) \cap F_f(\bar{x})$，得出矛盾.

参考文献：[61,93,128,146,150,203,238].

§4.4 基于择一定理的最优性条件

为导出问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D; \tag{1}$$

$$\min f(x), g(x) \leqslant 0, h(x) = 0; \tag{2}$$

及

$$\min f(x), g(x) \leqslant 0 \tag{3}$$

的最优性条件，一个基本的方法是运用适当的择一定理. 以下总设 $\bar{x}$ 是所考虑问题的可行点，$\tilde{f}(x)=f(x)-f(\bar{x})$，$f,g,h$ 在 $\bar{x}$ 可微.

为选用适当的择一定理，首先注意到以下基本事实：$\bar{x}$ 是(1)～(3)的最优解分别意味着以下问题无解：

$$\exists x\in D:\tilde{f}(x)<0,g(x)\leqslant 0,h(x)=0; \tag{4}$$

$$\exists x\in H:\tilde{f}(x)<0,g(x)\leqslant 0; \tag{5}$$

$$\exists x\in X:\tilde{f}(x)<0,g(x)\leqslant 0. \tag{6}$$

这就自然联系到备择的命题：

$$\exists(\rho,\lambda,\mu)\in(\mathbf{R}_+\times Y_+^*\times Z^*)\backslash\{0\}:$$
$$(\rho\tilde{f}+\lambda g+\mu h)(D)\geqslant 0; \tag{4'}$$

$$\exists\lambda\in Y_+^*:(\tilde{f}+\lambda g)(H)\geqslant 0; \tag{5'}$$

$$\exists\lambda\in Y_+^*:\tilde{f}+\lambda g\geqslant 0. \tag{6'}$$

若(4′)有解 $(\bar{\rho},\bar{\lambda},\bar{\mu})$，$L(\cdot,\rho,\lambda,\mu)=\rho f+\lambda g+\mu h$，则 $\bar{\rho}f(\bar{x})\leqslant L(D,\bar{\rho},\bar{\lambda},\bar{\mu})$，$0\leqslant\bar{\lambda}(g(\bar{x})\leqslant 0$，这推出 $\bar{\lambda}\in Q$，

$$\bar{\rho}f(\bar{x})=L(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})=\min L(D,\bar{\rho},\bar{\lambda},\bar{\mu}).$$

于是，当 D 凸时 $L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})\in(D-\bar{x})^*$ (1.6.3). 这就得到：

4.4.1 定理　设 $\bar{x}$ 是问题(1)的最优解. 若命题(4)与(4′)两择一且 D 凸，则有 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in\mathbf{R}_+\times Q\times Z^*$，使

$$L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})\in(D-x)^*.$$

4.4.1 可看作一个“Fritz John 型定理”而与 4.2.3 对照. 为将 4.4.1 中的“择一性条件”具体化，可应用某个 Motzkin 型定理，例如 3.4.2 或 3.6.8.

其次，若(5′)有解 $\bar{\lambda}$，则必定 $\bar{\lambda}\in Q$，且

$$f(\bar{x})=(f+\bar{\lambda}g)(x)=\min(f+\bar{\lambda}g)(H).$$

若 H 凸，则 $f'(\bar{x})+A^*\bar{\lambda}\in(H-\bar{x})^*$. 以上分析导致：

4.4.2 定理　设 Y_+^* 有弱* 紧凸基，$h(x)=Bx-b$，$R(B)$ 闭；(f,g) 在集 $H=h^{-1}(0)$ 上为 $\mathbf{R}_+\times Y_+$ - 次类凸，或为 * 拟凸且 * lsc；$\exists\hat{x}\in H:\langle Y_+^*\backslash\{0\},g(\hat{x})\rangle<0$. 若 $\bar{x}$ 是问题(2)的最优解，则 $\bar{x}$ 是 K-T 点.

证 在定理条件下，$(\bar{f},g)$ 亦必在 H 上为 $\mathbf{R}_+\times Y_+$-次类凸或为 * 拟凸且 * lsc，于是由定理条件及 3.6.6 推出命题(5′)有解 $\bar{\lambda}$. 前面已指出 $\bar{\lambda}\in Q$，$f'(\bar{x})+A^*\bar{\lambda}\in(H-\bar{x})^*$. 因 $H-\bar{x}=N(B)$，$N(B)^{\perp}=R(B^*)$(1.1.8)，故 $-f'(\bar{x})\in A^*Q+R(B^*)$，因此 $\bar{x}$ 是 K-T 点. □

取 $Z=\{0\}$ 从 4.4.2 得到如下推论.

4.4.3 推论 设 Y_+^* 有弱 * 紧凸基；(f,g) 为 $\mathbf{R}_+\times Y_+$-次类凸；$\exists\hat{x}\in X$：$\langle Y_+^*\backslash\{0\},g(\hat{x})\rangle<0$. 若 $\bar{x}$ 是问题(3)的最优解，则 $\bar{x}$ 是 K-T 点.

下面转向考虑局部最优性条件，为此需要用到切锥及一定的线性择一定理. 为简单起见，仅考虑问题(2)，(3). 若 $\bar{x}$ 是命题(2)(或(3))的局部最优解，则 $F_f(\bar{x})\cap K_M(\bar{x})=\varnothing$，$M=G\cap H$(或 $M=G$). 若 $N(B)\cap A^{-1}K_g\subset K_M(\bar{x})$(或 $A^{-1}K_g\subset K_G(\bar{x})$)，则 $F_f(\bar{x})\cap K_M(\bar{x})=\varnothing$ 推出

$$\exists x\in N(B):\langle f'(\bar{x}),x\rangle<0,Ax\in K_g \tag{7}$$

或

$$\exists x\in X:\langle f'(\bar{x}),x\rangle<0,Ax\in K_g \tag{8}$$

无解. 据此可以建立：

4.4.4 定理 设 K_g 与 $R(B)$ 为闭集，Q 有弱 * 紧凸基；$\exists\hat{x}\in N(B)$：$\langle Q\backslash\{0\},A\hat{x}\rangle<0$. 若 $\bar{x}$ 是问题(2)的局部最优解，$N(B)\cap A^{-1}K_g\subset K_M(\bar{x})$，则 $\bar{x}$ 是 K-T 点.

证 由定理条件推出(7)无解，K_g 是闭凸锥. 于是由 3.6.7 有 $\bar{\lambda}\in-K_g^*=Q$：$f'(\bar{x})+A^*\bar{\lambda}\in N(B)^{\perp}=R(B^*)$，从而 $\bar{x}$ 是 K-T 点. □

4.4.4 可与 4.3.4(iii) 对照. 4.4.4 中关键性的条件是 $N(B)\cap A^{-1}K_g\subset K_M(\bar{x})$. 当 $g,h\in C^1$，$(A,B):X\to Y\times Z$ 为满射时此条件满足(参看 4.1.1). 若 $Y=\mathbf{R}^m$，则 $Q=\mathbf{R}_+^l$，Q 有紧凸基是不成问题的.

取 $Z=\{0\}$ 从 4.4.4 得到：

4.4.5 推论 设 K_g 闭，$A^{-1}K_g\subset K_G(\bar{x})$，$Q$ 有弱 * 紧凸基；存

在 $\hat{x} \in X$，使 $\langle Q\backslash\{0\}, A\hat{x}\rangle < 0$. 若 $\bar{x}$ 是问题(3) 的局部最优解，则 $\bar{x}$ 是 K-T 点.

注 由 4.1.2，当 $g \in C^1$，$R(A) + K_g = Y$ 时 $A^{-1}K_g \subset K_0(\bar{x})$.

降低可微性条件，可建立以下定理.

4.4.6 定理 设 D 凸，$\mathring{Y}_+ \neq \varnothing$，当 $x \in D - \bar{x}$ 时 $D_+ f(\bar{x}, x)$ 与 $D_+ g(\bar{x}, x)$ 存在且 $(D_+ f(\bar{x}, \cdot), D_+ g(\bar{x}, \cdot))$ 在 $D - \bar{x}$ 上为 $\mathbf{R}_+ \times Y_+$ - 次类凸；$\bar{x}$ 是问题

$$\min f(x),\ g(x) \leqslant 0,\ x \in D$$

的局部最优解，则有 $0 \neq (\bar{\rho}, \bar{\lambda}) \in \mathbf{R}_+ \times Y_+^*$，使得

$$\bar{\rho} D_+ f(\bar{x}, x) + \bar{\lambda} D_+ g(\bar{x}, x) \geqslant -\bar{\lambda} g(\bar{x}), \quad x \in D - \bar{x}. \tag{9}$$

若 $D = X$ 且 $D_+ f(\bar{x}, \cdot)$ 与 $D_+ g(\bar{x}, \cdot)$ 是线性的，则 $\bar{\lambda} \in Q$，

$$\bar{\rho}\, D_+ f(\bar{x}, x) + \bar{\lambda}\, D_+ g(\bar{x}, x) = 0,\ x \in X.$$

证 $\bar{x}$ 是局部最优解推出以下问题无解：

$$\exists x \in D - \bar{x}: D_+ f(\bar{x}, x) < 0, D_+ g(\bar{x}, x) + g(\bar{x}) \ll 0. \tag{10}$$

事实上，若 x 是式(10) 的解，则当 $t > 0$ 充分小时，

$$t^{-1}\Delta g(\bar{x}, tx) + g(\bar{x}) \leqslant 0,$$

因此 $g(\bar{x} + tx) \leqslant (1 - t)g(\bar{x}) \leqslant 0$，$\bar{x} + tx \in D$. 于是 $f(\bar{x} + tx) \geqslant f(\bar{x})$，这推出 $D_+ f(x, x) \geqslant 0$，得出矛盾. 故可用 3.2.2 得出定理结论. □

4.4.6 是[85, Th. 4.1]与[104, Th. 4.1] 的一个改进. 关于类似于 4.4.6 的结果还可参看[104, Th. 5.2].

§4.5 充分条件

考虑最优化问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D. \tag{1}$$

设 $\bar{x} \in M = D \cap G \cap H$，$(\bar{\rho},\bar{\lambda},\bar{\mu}) \in \mathbf{R}_+ \times Q \times Z^*$ 满足

$$L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu}) = \bar{\rho}\Gamma + A^*\bar{\lambda} + B^*\bar{\mu} = 0. \tag{2}$$

本节给出 $\bar{x}$ 为最优解的某些充分条件，这类条件通常涉及 f,g,h（或其组合）的某种凸性，与条件(2)一起合称为“John 充分条件”，当 $\bar{\rho} = 1$ 时称为“Kuhn-Tucker 充分条件”. 以下提到的 $F: M \times X^* \to \mathbf{R}$ 与 $\eta: M \to X$ 是任给的，在同一命题中保持固定，$F(x,\cdot)$ 是次线性的(参看 1.5.5).

4.5.1 定理 设 $(\bar{\rho},\bar{\lambda},\bar{\mu}) \in \mathbf{R}_+ \times Q \times Z^*$ 满足(2)，则以下每个条件推出 $\bar{x}$ 为问题(1)的严格最优解：

(i) $L(\cdot,\bar{\rho},\bar{\lambda},\bar{\mu})$ 在 $\bar{x}$ 关于 M 为 (F,r)- 严格伪凸，$r \geqslant 0$；

(ii) $\bar{\rho}f$ 与 $\bar{\lambda}g + \bar{\mu}h$ 之一在 $\bar{x}$ 关于 M 为 (F,r)- 严格伪凸，另一个为 (F,r')- 拟凸，$r + r' \geqslant 0$；

(iii) $\bar{\rho}f,\bar{\lambda}g,\bar{\mu}h$ 之一在 $\bar{x}$ 关于 M 为 (F,r)- 严格伪凸，另两个分别为 (F,r')- 拟凸与 (F,r'')- 拟凸，$r + r' + r'' \geqslant 0$.

若将以上条件中的“(F,r)-”等皆换成“η-”，则结论仍成立.

证 不妨设 $\bar{\rho}f,\bar{\lambda}g,\bar{\mu}h$ 在 $\bar{x}$ 关于 M 分别为 (F,r)- 严格伪凸、(F,r')- 拟凸与 (F,r'')- 拟凸，其余情况的证明是类似的. 若 $\bar{x}$ 非严格最优解，则有 $x \in M\backslash\{\bar{x}\}$：$f(x) \leqslant f(\bar{x})$，从而 $\bar{\rho}f(x) \leqslant \bar{\rho}f(\bar{x})$. 于是，由 $\bar{\rho}f$ 在 $\bar{x}$ 为 (F,r)- 严格伪凸有

$$F(x,\bar{\rho}\Gamma) + r|x - \bar{x}| < 0.$$

其次，由 $\bar{\lambda}g(x) \leqslant 0 = \bar{\lambda}g(\bar{x})$ 及 $\bar{\lambda}g$ 的 (F,r')- 拟凸性有

$$F(x,A^*\bar{\lambda}) + r'|x - \bar{x}| \leqslant 0.$$

同理

$$F(x,B^*\bar{\mu}) + r''|x - \bar{x}| \leqslant 0.$$

综合以上三式及(2)得

$$\begin{aligned} 0 &= F(x,\bar{\rho}\Gamma + A^*\lambda + B^*\bar{\mu}) \\ &\leqslant F(x,\bar{\rho}\Gamma) + F(x,A^*\bar{\lambda}) + F(x,B^*\bar{\mu}) \\ &< -(r + r' + r'')|x - \bar{x}| \leqslant 0, \end{aligned}$$

得出矛盾. □

4.5.1 综合并推广了[5,6,83,120,171]中的一些结果.

若 $\forall x \in M\backslash\{\bar{x}\}: g(x) \leqslant g(\bar{x}) \Rightarrow g'(\bar{x})(x-\bar{x}) \ll 0$，则说 g 在 $\bar{x}$ 关于 M 为**强伪凸**. 当 $Y = \mathbf{R}$ 时，强伪凸即严格伪凸.

4.5.2 定理 设(2)满足且 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+ \times Q \times Z^*$；$f$ 在 $\bar{x}$ 关于 M 为伪凸[严格伪凸]，g, h 在 $\bar{x}$ 关于 M 皆为强伪凸，$g(\bar{x}) = 0$. 则 $\bar{x}$ 是问题(1)的最优解[严格最优解].

证 若结论不真，则 $\exists x \in M\backslash\{\bar{x}\}: f(x) < f(\bar{x})[f(x) \leqslant f(\bar{x})]$，从而 $Tz < 0, z = x - \bar{x}$. 其次，由 $g(x) \leqslant 0 = g(\bar{x})$ 及 g 强伪凸有 $Az \ll 0$；同理 $Bz \ll 0$. 于是

$$
\begin{aligned}
0 &= \langle \bar{\rho} T + A^* \bar{\lambda} + B^* \bar{\mu}, z \rangle \\
&= \langle (\bar{\rho}, \bar{\lambda}, \bar{\mu}), (Tz, Az, Bz) \rangle < 0,
\end{aligned}
$$

得出矛盾. □

以下设 $D = X, L(\cdot, \lambda, \mu) = f + \lambda g + \mu h$. 用证4.5.1的方法容易证明：

4.5.3 定理 设 $\bar{x}$ 是问题(1)的K-T点，$\bar{\lambda}, \bar{\mu}$ 是K-T乘子. 则以下每个条件推出 $\bar{x}$ 为问题(1)的最优解：

(i) $L(\cdot, \bar{\lambda}, \bar{\mu})$ 在 $\bar{x}$ 关于 M 为 (F, r)-伪凸，$r \geqslant 0$；

(ii) f 与 $\bar{\lambda}g + \bar{\mu}h$ 在 $\bar{x}$ 关于 M 分别为 (F, r)-伪凸与 (F, r')-拟凸，$r + r' \geqslant 0$；

(iii) $f, \bar{\lambda}g$ 与 $\bar{\mu}h$ 在 $\bar{x}$ 关于 M 分别为 (F, r)-伪凸、(F, r')-拟凸与 (F, r'')-拟凸，$r + r' + r'' \geqslant 0$.

4.5.3 综合并推广了[8, 83, 120]中的一些结果.

由4.5.3推出，若 $f, \bar{\lambda}g$ 与 $\bar{\mu}h$ 在 $\bar{x}$ 关于 M 皆为 η-凸，则 $\bar{x}$ 是(1)的最优解. 令 $\varphi = (f, \bar{\lambda}g, \bar{\mu}h): X \to \mathbf{R}^3$，则 $f, \bar{\lambda}g, \bar{\mu}h$ 在 $\bar{x}$ 关于 M 为 η-凸相当于

$$\forall x \in M, \exists \eta(x) \in X: \varphi'(\bar{x})\eta(x) \leqslant \varphi(x) - \varphi(\bar{x}). \tag{3}$$

由3.3.6，条件(3)等价于：$\forall x \in M$，以下问题无解：

$$\exists \alpha \in \mathbf{R}^3_+: \varphi'(\bar{x})^* \alpha = 0, \langle \alpha, \varphi(\bar{x}) - \varphi(x) \rangle = 1. \tag{4}$$

令 $\alpha = (\alpha_i)$，可将(4)中的方程写成：

$$\begin{cases}\alpha_1 f'(\bar{x}) + \alpha_2 A^* \bar{\lambda} + \alpha_3 B^* \bar{\mu} = 0, \\ \alpha_1[f(x) - f(\bar{x})] + \alpha_2 \langle \bar{\lambda}, g(x) \rangle = -1. \end{cases} \tag{5}$$

于是得到：

4.5.4 推论 若 $\forall x \in M$，方程组(5)不存在解 $\alpha_i \geqslant 0 (1 \leqslant i \leqslant 3)$，则 $\bar{x}$ 是问题(1)的最优解.

[83] 研究了以下例子

$$\min f(x_1, x_2),\ g(x_1, x_2) \leqslant 0, \tag{6}$$

其中 $f = x_1 - \sin x_2, g = (g_i) : \mathbf{R}^2 \to \mathbf{R}^6$,

$$g_1 = \sin x_1 - 4 \sin x_2,$$
$$g_2 = 2 \sin x_1 + 7\sin x_2 + x_1 - 6,$$
$$g_3 = 2(x_1 + x_2) - 3,$$
$$g_4 = 4(x_1^2 + x_2^2) - 9,$$
$$g_5 = -\sin x_1,$$
$$g_6 = -\sin x_2.$$

可验知 $\bar{x} = (0, \arcsin(6/7))$ 为 K-T 点，相应的 K-T 乘子为

$$\bar{\lambda} = (0, 1/7, 0, 0, 10/7, 0).$$

从(5)的第一个方程得出 $\alpha_1 = \alpha_2$；代入第二个方程得

$$8\alpha_1(x_1 - \sin x_1) = \quad 7.$$

任何 $\alpha_1 \geqslant 0$ 不满足以上方程. 因此，由 4.5.4 知 $\bar{x}$ 是(6)的最优解.

实际上，取 $\eta = (\sin x_1, (7\sin x_2 - 6)/\sqrt{13})$，可验知 $f, g_i (1 \leqslant i \leqslant 6)$ 在 $\bar{x}$ 皆为 η-凸.

下面考虑的一类充分条件不明显涉及 K-T 乘子.

4.5.5 定理 设 $\dim X < \infty$，以下条件之一满足：

(i) $-f'(\bar{x}) \in (A^*Q + R(B^*))^\circ$；

(ii) K_g 闭且 $\langle f'(\bar{x}), (N(B) \cap A^{-1}K_g) \setminus \{0\} \rangle > 0$.

则 $\exists \varepsilon, \delta > 0, \forall x \in M \cap B_\delta(\bar{x}) : f(x) \geqslant f(\bar{x}) + \varepsilon |x - \bar{x}|$.

从而 $\bar{x}$ 是问题(1)的严格局部最优解.

证 设定理结论不成立，则 $\forall n \geqslant 1, \exists x_n \in M \cap B(\bar{x}, 1/n)$：

$f(x_n) < f(\overline{x}) + n^{-1}|x_n - \overline{x}|$. 必定 $t_n \stackrel{\text{def}}{=} |x_n - \overline{x}| > 0$. 令 $z_n = (x_n - \overline{x})/t_n$，则 $x_n = \overline{x} + t_n z_n$，$|z_n| = 1$. 不妨设 $z_n \to z \neq 0$($\dim X < \infty$ 用于此!). 由

$$f(x_n) - f(\overline{x}) = t_n\langle f'(\overline{x}), z_n\rangle + o(t_n) < t_n/n$$

推出 $\langle f'(\overline{x}), z\rangle \leqslant 0$. 由

$$0 \geqslant \langle Q, g(x_n) - g(\overline{x})\rangle = \langle Q, t_n A z_n + o(t_n)\rangle$$

推出 $\langle Q, Az\rangle \leqslant 0$；类似地有 $Bz = 0$. 于是

$$\langle A^* Q + R(B^*), z\rangle = \langle Q, Az\rangle + \langle Z^*, Bz\rangle \leqslant 0,$$

因此 $-z \in (A^*Q + R(B^*))^*$. 若条件(i)满足，则用 1.4.2(ii) 得 $\langle f'(\overline{x}), z\rangle > 0$，引出矛盾. 若条件(ii)满足，则 $K \stackrel{\text{def}}{=} N(B) \cap A^{-1}K_g$ 是 X 中的闭凸锥，$\langle f'(\overline{x}), K\backslash\{0\}\rangle > 0$，于是 $f'(\overline{x}) \in \operatorname{int} K^*$(1.4.4)；再结合1.4.7，3.1.2 得

$$\begin{aligned} -f'(\overline{x}) &\in -[N(B)^\perp + (A^{-1}K_g)^*] \\ &= (R(B^*) - A^*K_g^*)^\circ \\ &= (R(B^*) + A^*Q)^\circ. \end{aligned}$$

这表明条件(i)满足，由已证结论知此为不可能. □

4.5.5 推广了[150, Th. 5.1]. 注意 4.5.5 中的条件 $\dim X < \infty$ 是重要的，试看以下反例.

4.5.6 例([150]) 设 $X = Y = l^2$，$Y_+ = \{(x_n) \in l^2 : x_n \geqslant 0(\forall n)\}$；$f(x) = (a, x) - |x|^2$，$a = (1/n)_{n\geqslant 1}$，$g(x) = -x$. 考虑

$$\min f(x),\ g(x) \leqslant 0. \tag{7}$$

取 $\overline{x} = 0$，则 $f'(\overline{x}) = a$，$A = g'(\overline{x}) = -I$，$g(\overline{x}) = 0$，$K_g = Y_-$，$A^{-1}K_g = Y_+$，$\langle f'(\overline{x}), (A^{-1}K_g)\backslash\{0\}\rangle = \langle a, Y_+\backslash\{0\}\rangle > 0$，这表明 4.5.5 之条件(ii)满足. 但 $\overline{x} = 0$ 不是问题(7)的局部最优解：令 $x^{(n)} = (0, \cdots, 0, 2/n, 0, \cdots)$，则 $x^{(n)} \to 0(n \to \infty)$，$g(x^{(n)}) \leqslant 0$，$f(x^{(n)}) = -2/n^2 < f(0)$!

为了对无限维的 X 建立类似结果，引进概念：

4.5.7 定义 设 $\overline{x} \in D \subset X$，$K \subset X$ 是某个锥. 若存在映射

$\varphi: D \to K$,使得 $\varphi(x) = x - \bar{x} + o(|x - \bar{x}|)(x \in D)$,则说 K 在 $\bar{x}$ 逼近 D.

4.5.8 定理 假定:(i) 存在锥 $K \subset X$ 在 $\bar{x}$ 逼近 M,(ii) $\exists \beta > 0, \forall z \in K: \langle f'(\bar{x}), z \rangle \geqslant \beta |z|$,则 4.5.5 的结论成立.

证 取 $\varphi: M \to K$,使得 $\varphi(x) = x - \bar{x} + o(|x - \bar{x}|)(x \in M)$,则当 $x \in M$ 且 $x \to \bar{x}$ 时,有

$$\begin{aligned} f(x) - f(\bar{x}) &= \langle f'(\bar{x}), x - \bar{x} \rangle + o(|x - \bar{x}|) \\ &= \langle f'(\bar{x}), \varphi(x) \rangle + o(|x - \bar{x}|) \\ &\geqslant \beta |\varphi(x)| + o(|x - \bar{x}|) \\ &= \beta |x - \bar{x}| + o(|x - \bar{x}|) \\ &\geqslant (\beta/2) |x - \bar{x}|, \end{aligned}$$

这表明 4.5.5 的结论成立. □

4.5.8 蕴涵了[150,Th. 5.3]. 注意 4.5.8 完全未涉及锥 K 的结构. 不过,"K 在 $\bar{x}$ 逼近 M"这一条件使得 K 并无多大选择余地. 4.1.1 启示出,可取 $K = N(B) \cap A^{-1}K_g$.

4.5.9 命题 以下每个条件推出 $K = N(B) \cap A^{-1}K_g$ 在 $\bar{x}$ 逼近 M:

(i) $R(A, B) + K_g \times 0 = Y \times Z$;

(ii) $\dim X < \infty$ 且 K_g 闭.

证 关于条件(i)的证明参看[150,Th. 4.2]. 下面设条件(ii)满足. 对每个 $x \in M$,取 $\varphi(x) \in K$,使

$$|\varphi(x) - (x - \bar{x})| = d(x - \bar{x}, K \cap B(0, 2|x - \bar{x}|)). \quad (8)$$

若 $\varphi(x) - (x - \bar{x}) \neq o(|x - \bar{x}|)$,则有 $\varepsilon > 0, x_n \in M, x_n \to \bar{x}$,使得 $|\varphi(x_n) - (x_n - \bar{x})| > \varepsilon |x_n - \bar{x}|$. 令 $t_n = |x_n - \bar{x}|, z_n = (x_n - \bar{x})/t_n$,可设 $z_n \to z \neq 0$,则 $z \in K_M(x) \subset K$(4.1.1). 由 $t_n z \in K$, $|t_n z| = t_n < 2t_n = 2|x_n - \bar{x}|$ 及(8)推出

$$\begin{aligned} \varepsilon t_n &< |\varphi(x_n) - (x_n - \bar{x})| \\ &\leqslant |x_n - \bar{x} - t_n z| = t_n |z_n - z|, \end{aligned}$$

这与 $z_n \to z$ 矛盾. 因此 K 在 $\bar{x}$ 逼近 M. □

下面的结果使得有可能减弱 4.5.8 中的条件(ii).

4.5.10 命题 若锥 $K \subset X$ 有弱紧基 D,则 4.5.8 中的条件(ii) 可代以条件 $\langle f'(\bar{x}), K\backslash\{0\}\rangle > 0$.

证 设 4.5.8 中条件(ii) 不满足,则 $\forall n \geq 1$, $\exists x_n \in K$: $\langle f'(\bar{x}), x_n\rangle < |x_n|/n$. 设 $x_n = t_n z_n, z_n \in D$, 则 $\langle f'(\bar{x}), z_n\rangle < |z_n|/n$. 不妨设 $z_n \rightharpoonup z \in D$,因此 $\{z_n\}$ 有界,故 $\langle f'(\bar{x}), z\rangle \leqslant 0$,条件 $\langle f'(\bar{x}), K\backslash\{0\}\rangle > 0$ 不能满足. □

结合 4.5.8～4.5.10 推出,4.5.5 的条件(ii) 对于 $\bar{x}$ 为严格局部最优解是充分的.

参考文献:[5,6,8,83,84,120,149,150,165,171,183].

§4.6 非光滑最优性条件

本节对问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D \tag{1}$$

建立 §4.2～§4.5 中部分结果的非光滑推广. 设 $\bar{x} \in M = D \cap G \cap H$, f, g, h 在 $\bar{x}$ 邻近为 Lip.

4.6.1 定理 设 $Y = \mathbf{R}^m$, $K \subset F_D(\bar{x})$ 是一非空开凸锥;h 在 $\bar{x}$ 邻近为 C^1 映射,$R(B)$ 为闭集. 若 $\bar{x}$ 是问题(1) 的局部最优解,则有 $0 \neq (\bar{p}, \bar{\lambda}, \mu) \in \mathbf{R}_+ \times \mathbf{R}_+^l \times Z^*$, $\bar{\lambda} = (\bar{\lambda}_i)$ (以下概如此),使得

$$\bar{p}\partial f(\bar{x}) + \sum \bar{\lambda}_i \partial g_i(\bar{x}) + B^* \bar{\mu}) \cap K^* \neq \varnothing. \tag{2}$$

若 D 凸且 $D^\circ \neq \varnothing$,则可取 $K = F_D(\bar{x})$, $K^* = (D - \bar{x})^*$;若 $\bar{x} \in D^\circ$, $K = F_D(\bar{x})$, 则(2)成为

$$0 \in \bar{p}\partial f(\bar{x}) + \sum \bar{\lambda}_i \partial g_i(\bar{x}) + B^* \bar{\mu}. \tag{3}$$

证 不妨设 $g(\bar{x}) = 0$(否则以 g_I 取代 g),$0 \bar{\in} \partial f(\bar{x}) \cup \partial g_i(\bar{x})$ $(1 \leqslant i \leqslant m)$, $R(B) = Z$(否则(3) 式自动成立). 令 $K_0 = \{z : f^\circ(\bar{x}, z) < 0\}$, $K_i = \{z : g_i^\circ(\bar{x}, z) < 0\}$, $K_{m+1} = K$, $K_n = N(B) = K_H(\bar{x})$,

$n=m+2$. 由 $0\overline{\in}\partial f(\bar{x})$ 推出 $K_0\neq\varnothing$，因此 K_0 是非空开凸锥且 $K_0^*=\mathbf{R}_-\partial f(\bar{x})$(4.1.9). $K_i(1\leqslant i\leqslant m)$ 仿此. 若 $z\in X\backslash F_f(\bar{x})$，则 $\exists t_k\downarrow 0, z_k\to z, x_k=\bar{x}+t_kz_k$，使 $f(x_k)\geqslant f(\bar{x})$. 于是

$$f^\circ(\bar{x},z)\geqslant\limsup_k\frac{f(\bar{x}+t_kz)-f(\bar{x})}{t_k}$$

$$\geqslant\limsup_k\frac{f(\bar{x}+t_kz)-f(x_k)}{t_k}=0,$$

故 $z\overline{\in}K_0$. 这就证得 $K_0\subset F_f(\bar{x})$. 令 $G_i=g_i^{-1}(\mathbf{R}_-)$，则 $G=\bigcap G_i$，$K_i\subset F(G_i,\bar{x})$(4.1.9)，从而 $\bigcap_1^m K_i\subset\bigcap F(G_i,\bar{x})=F_G(\bar{x})$(2.5.2(i)). 于是

$$\bigcap_{i=0}^n K_i\subset F_f(\bar{x})\cap F_G(\bar{x})\cap F_D(\bar{x})\cap K_H(\bar{x})$$

$$\subset F_f(\bar{x})\cap K_M(\bar{x})=\varnothing.$$

由 4.2.2，有 $0\neq(u_i)\in\coprod_0^n K_i^*$，使 $\sum u_i=0$. 设 $u_0=-\bar{\rho}\,\bar{u}_0, u_i=-\bar{\lambda}_i\bar{u}_i, \bar{\rho},\bar{\lambda}_i\geqslant 0, \bar{u}_0\in\partial f(\bar{x}), \bar{u}_i\in\partial g_i(\bar{x})(1\leqslant i\leqslant m), \bar{\lambda}=(\bar{\lambda}_i)$, $u_n=-B^*\bar{\mu}, \bar{\mu}\in Z^*$，则 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in\mathbf{R}_+\times\mathbf{R}_+^m\times Z^*$，

$$u_{m+1}=-u_0-\sum_{i=1}^m u_i-u_n$$

$$\in[\bar{\rho}\partial f(\bar{x})+\sum\bar{\lambda}_i\partial g_i(\bar{x})+B^*\mu]\cap K^*,$$

这表明(2)成立. 其余结论是明显的. □

4.6.1 可看作 4.2.4 的一个非光滑推广. 若 $Z=\mathbf{R}^p$，则 4.2.4 有如下更完全的推广.

4.6.2 定理 设 $Y=\mathbf{R}^m, Z=\mathbf{R}^p$，$D$ 为闭集. 若 $\bar{x}$ 是问题(1)的局部最优解，$k>0$ 充分大，则有 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in\mathbf{R}_+\times\mathbf{R}_+^l\times\mathbf{R}^p$，使得

$$0\in\partial_x L(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})+k\partial d_D(\bar{x}). \tag{4}$$

若 $\bar{x}\in D^\circ$，则(4) 成为

$$0\in\partial_x L(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu}). \tag{5}$$

证 不妨设 $f(\bar{x})=0, g(\bar{x})=0$. 取充分小的 $\varepsilon>0$，令 $f_\varepsilon=$

$f+\varepsilon, \omega_\varepsilon=(f_\varepsilon, g, h), T=(\mathbf{R}_+\times\mathbf{R}_+^m\times\mathbf{R}^p)\cap S^{m+p}$($S^k$ 记 k 维单位球面). $\forall t=(\rho,\lambda,\mu)\in T$, 定义

$$\varphi_{t\varepsilon}(x)=\langle t,\omega_\varepsilon(x)\rangle=L(x,t)+\rho\varepsilon;$$

$\varphi_\varepsilon(x)=\max\limits_{t\in T}\varphi_{t\varepsilon}(x)$. 因问题是局部的,不妨设 Lip $\omega_\varepsilon<k$,这就推出 Lip $\varphi_{t\varepsilon}<k$,从而 Lip $\varphi_\varepsilon<k$.

任给 $x\in D$,有 $t=(\rho,\lambda,\mu)\in T,\lambda=(\lambda_i),\mu=(\mu_j)$, 使得

$$\varphi_\varepsilon(x)=\varphi_{t\varepsilon}(x)=\rho f_\varepsilon(x)+\sum\lambda_i g_i(x)+\sum\mu_j h_j(x). \quad (6)$$

因 $\bar{x}$ 是最优解,容易看出式(6) 右端 $1+m+p$ 项皆非负. 因此,$\rho\neq 0\to f_\varepsilon(x)\geqslant 0;\lambda_i\neq 0\Rightarrow g_i(x)\geqslant 0$. 令

$$\Phi_\varepsilon=(f_\varepsilon^+,g_1^+,\cdots,g_m^+,h)\ (f^+=\max\{f,0\}),$$

则可指明

$$\varphi_\varepsilon(x)=\langle t,\Phi_\varepsilon(x)\rangle=|\Phi_\varepsilon(x)|,$$

$t=\Phi_\varepsilon(x)/|\Phi_\varepsilon(x)|$ 由 x 唯一确定.

直接看出 $\varphi_\varepsilon(\bar{x})=\varepsilon\leqslant\inf\limits_D\varphi_\varepsilon(x)+\varepsilon$. 取 $\delta=\sqrt{\varepsilon}$,由 1.6.1 有 $x_\varepsilon\in D\cap\overline{B}(\bar{x},\sqrt{\varepsilon})$, 使得

$$\varphi_\varepsilon(x_\varepsilon)=\min_{x\in D}[\varphi_\varepsilon(x)+\sqrt{\varepsilon}\,|x-x_\varepsilon|]. \quad (7)$$

因 ε 充分小,可设 Lip $(\varphi_\varepsilon+\sqrt{\varepsilon}\,|\cdot-x_\varepsilon|)\leqslant k$,于是由(7),2.2.9 及 2.1.2 之 3° 得

$$0\in\partial(\varphi_\varepsilon+kd_D)(x_\varepsilon)+\sqrt{\varepsilon}\,B^*, \quad (8)$$

其中 $B^*=\overline{B}_1(0)\subset X^*$. 令 $R_{t\varepsilon}=\varphi_{t\varepsilon}+kd_D, S_t=L(\cdot,t)+kd_D$, $R_\varepsilon=\varphi_\varepsilon+kd_D$,则Lip $(\varphi_{t\varepsilon}-\varphi_{\tau\varepsilon})=$ Lip $(R_{t\varepsilon}-R_{\tau\varepsilon})\leqslant k|t-\tau|$,于是 $\partial R_{t\varepsilon}(x)\subset\partial R_{\tau\varepsilon}(x)+k|t-\tau|B^*$ (2.2.2(i)). 这结合2.2.2(ii)得出 $\partial_T R_{t\varepsilon}(x)=\partial R_{t\varepsilon}(x)$,记号依 2.4.1. 对 S_t 有类似结论. 因 $R_\varepsilon=\max\limits_T R_{t\varepsilon},\partial R_{t\varepsilon}=\partial S_t$,故由 2.4.1 与式(8) 有 $0\in\partial S_{t_\varepsilon}(x_\varepsilon)+\sqrt{\varepsilon}\,B^*$, t_ε 由 x_ε 唯一决定.

取 $\varepsilon_n\downarrow 0$,则 $x_n\overset{\text{def}}{=}x_{\varepsilon_n}\to\bar{x}$,且不妨设 $t_n\overset{\text{def}}{=}t_{\varepsilon_n}\to\bar{t}=(\bar{\rho},\bar{\lambda},\bar{\mu})\in$

T. 于是

$$0 \in \partial S_t(x_n) + (k|t_n - t| + \sqrt{\varepsilon_n})B^*,$$

这结合 2.2.2(ii)得 $0 \in \partial S_t(\bar{x})$,由此得出(4). □

注 利用 2.3.3 从式(4)推出

$$0 \in \bar{\rho}\partial f(\bar{x}) + \sum \bar{\lambda}_i \partial g_i(\bar{x}) + \partial(\bar{\mu}h)(\bar{x}) + k\partial d_D(\bar{x}).$$

另一方面,设 D 凸且 $D^\circ \neq \varnothing$,在 4.6.1 中取 $K = F_D(\bar{x})$,则 $K^* = (D - \bar{x})^* = -\partial\delta_D(\bar{x}) \subset \mathbf{R}\ \ \partial d_D(\bar{x})$(2.1.2,2.5.6),于是由式(2)推出

$$0 \in \bar{\rho}\partial f(\bar{x}) + \sum \bar{\lambda}_i \partial g_i(\bar{x}) + \partial(\bar{\mu}h)(\bar{x}) + \mathbf{R}_+ \partial d_D(\bar{x}).$$

对比以上两式,可看出 4.6.1 与 4.6.2 是高度类似的.

现在转向推广 § 4.3 中的结果. 为行文方便起见,引入以下术语(对照 4.3.1).

4.6.3 定义 若存在 $(\bar{\lambda}, \bar{\mu}) \in Q \times Z^*$,使得

$$0 \in \partial f(\bar{x}) + \partial g_{\bar{\lambda}}(\bar{x}) + \partial h_{\bar{\mu}}(\bar{x}) \tag{9}$$

(其中 $g_{\bar{\lambda}} = \bar{\lambda}g, h_{\bar{\mu}} = \bar{\mu}h$),或当 $Y = \mathbf{R}^m, Z = \mathbf{R}^p$ 时有

$$0 \in \partial f(\bar{x}) + \sum \bar{\lambda}_i \partial g_i(\bar{x}) + \sum \bar{\mu}_j \partial h_j(\bar{x}), \tag{10}$$

则称 $\bar{x}$ 为问题(1)的**广义 K-T 点**,而称 $\bar{\lambda}, \bar{\mu}$(或 $\bar{\lambda}_i, \bar{\mu}_j$)为相应的广义 K-T 乘子.

若(3)或(5)满足且其中 $\bar{\rho} \neq 0$,则 $\bar{x}$ 必为广义 K-T 点. 因此,基于 4.6.1 与 4.6.2 可建立:

4.6.4 定理 设 $\bar{x} \in D^\circ$ 是问题(1)的局部最优解,$Y = \mathbf{R}^m$,则以下每个条件推出 $\bar{x}$ 是(1)的广义 K-T 点:

(i) h 在 $\bar{x}$ 邻近为 C^1 映射,$R(B) = Z$,且存在 $z \in N(B)$,使得 $g_i^\circ(x, z) < 0\ (\forall i \in I)$;

(ii) $g_i(\forall i \in I)$ 在 $\bar{x}$ 关于 M 为广义伪凸,$h(x) = Bx - b$,$R(B) = Z$,且 $\exists \hat{x} \in h^{-1}(0)$;$g_I(\hat{x}) \ll 0$;

(iii) $Z = \mathbf{R}^p$,$\exists z \in X$:$g_i^\circ(\bar{x}, z) < 0\ (\forall i \in I)$,$h_j^\circ(\bar{x}, \pm z) =$

$0, \forall u_j \in \partial h_j(\bar{x})$: $\{u_j : 1 \leqslant j \leqslant p\}$ 线性无关；

(iv) $Z = \mathbf{R}^p, \forall u_i \in \partial g_i(\bar{x}), \forall v_j \in \partial h_j(\bar{x})$: $\{u_i, v_j : i \in I, 1 \leqslant j \leqslant p\}$ 线性无关.

证 假定条件(i)～(iv)之一满足，则依 4.6.1 或 4.6.2 有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+ \times \mathbf{R}_+^I \times Z^*$，使式(3)或式(5)成立. 设 $\bar{\rho} = 0$，则 $(\bar{\lambda}, \bar{\mu}) \neq 0$，

$$0 \in \sum \bar{\lambda}_i \partial g_i(\bar{x}) + \partial h_{\bar{\mu}}(\bar{x}). \tag{11}$$

若条件(i)满足，则必 $\bar{\lambda} \neq 0$. 否则 $\bar{\mu} \neq 0$，但由式(11)有 $B^* \bar{\mu} = 0$，矛盾于 $R(B) = Z$. 设 z 如条件(i)，则

$$\begin{aligned} \langle \sum \bar{\lambda}_i \partial g_i(\bar{x}) + B^* \bar{\mu}, z \rangle &= \sum \bar{\lambda}_i \langle \partial g_i(\bar{x}), z \rangle \\ &\leqslant \sum \bar{\lambda}_i g_i^\circ(\bar{x}, z) < 0, \end{aligned}$$

这矛盾于(11). 若条件(ii)满足，则亦有 $\bar{\lambda} \neq 0$. 令 $z = \hat{x} - \bar{x}$，则 $Bz = h(\hat{x}) - h(\bar{x}) = 0$. $\forall i \in I$，由 $g_i(\hat{x}) < 0 = g_i(\bar{x})$ 及 g_i 在 $\bar{x}$ 关于 M 为广义伪凸推出 $g_i^\circ(\bar{x}, z) < 0$（§2.2(10)），可见条件(i)必满足，因而亦导致矛盾. 若条件(iii)满足，则亦有 $\bar{\lambda} \neq 0$. 否则 $\bar{\mu} \neq 0$，由(11)有 $0 \in \sum \bar{\mu}_j \partial h_j(\bar{x})$，矛盾于条件(iii)中的线性无关假设. 设 z 如条件(iii)，则

$$\begin{aligned} &\langle \sum \bar{\lambda}_i \partial g_i(\bar{x}) + \sum \bar{\mu}_j \partial h_j(\bar{x}),\ z \rangle \\ &= \sum \bar{\lambda}_i \langle \partial g_i(\bar{x}),\ z \rangle + \sum \bar{\mu} \langle \partial h_j(\bar{x}),\ z \rangle \\ &\leqslant \sum \bar{\lambda}_i g_i^\circ(\bar{x},\ z) + \sum |\bar{\mu}_j| h_j^\circ(\bar{x},\ z\, \mathrm{sgn}\mu_j) < 0, \end{aligned}$$

亦矛盾于(11). 条件(iv)显然与(11)不相容. □

4.6.4 可看作 4.3.3 的非光滑推广.

为推广 4.3.4，使用如下的集（参看 §4.1(4)，(5)）.

$$K = \{z : g_\lambda^\circ(\bar{x}, z) \leqslant 0, h_\mu^\circ(\bar{x}, z) \leqslant 0 (\forall \lambda \in Q, \mu \in Z^*)\}; \tag{12}$$

$$\Phi = \bigcup \{\partial g_\lambda(\bar{x}) + \partial h_\mu(\bar{x}) : (\lambda, \mu) \in Q \times Z^*\}. \tag{13}$$

显然，$\bar{x}$是(1)的广义K-T点 $\Leftrightarrow -\partial f(\bar{x})\cap\Phi\neq\varnothing$. 基于此，可以建立：

4.6.5 定理 设$\bar{x}$是问题(1)的局部最优解；Φ弱*闭凸且$K\subset K_M(\bar{x})$，则$\bar{x}$是(1)的广义K-T点.

证 若$\bar{x}$不是广义K-T点，则$-\partial f(\bar{x})\cap\Phi=\varnothing$. 由1.1.3有$z\in X$，使$\langle-\partial f(\bar{x}),z\rangle>\langle\Phi,z\rangle$. 这一方面推出$\langle\partial f(\bar{x}),z\rangle<0$，从而$f^\circ(\bar{x},z)<0$，$z\in F_f(\bar{x})$(参考4.6.1之证)；另一方面推出，$\forall(\lambda,\mu)\in Q\times Z^*$，有$\langle\partial g_\lambda(x),z\rangle\leqslant 0$，$\langle\partial h_\mu(\bar{x}),z\rangle\leqslant 0$，从而$g_\lambda^\circ(\bar{x},z)\leqslant 0$，$h_\mu^\circ(\bar{x},z)\leqslant 0$. 这表明$z\in K\subset K_M(\bar{x})$. 但这与$F_f(\bar{x})\cap K_M(\bar{x})=\varnothing$(4.2.1)矛盾. □

注 在4.1.6的条件下，有$K_M(\bar{x})\subset K$，此时4.6.5中的条件$K\subset K_M(\bar{x})$相当于$K=K_M(\bar{x})$. 这可看作某种"约束品性".

4.6.6 引理 若$\bar{x}$是问题(1)的局部最优解，则

$$\partial f(\bar{x})\cap T_M^*(\bar{x})\neq\varnothing.$$

证 若$\partial f(\bar{x})\cap T_M^*(\bar{x})=\varnothing$，则由1.1.3有$z\in X$，使得$\langle\partial f(\bar{x}),z\rangle<\langle T_M^*(\bar{x}),z\rangle$. 这推出$\langle\partial f(\bar{x}),z\rangle<0$，从而$f^\circ(\bar{x},z)<0$，$z\in F_f(\bar{x})$；$\langle T_M^*(\bar{x}),z\rangle\geqslant 0$，从而$z\in T_M(\bar{x})\subset K_M(\bar{x})$(注意$T_M(\bar{x})$是闭凸锥). 这矛盾于$F_f(\bar{x})\cap T_M(\bar{x})=\varnothing$. □

利用4.6.6，易证得类似于4.6.5的以下结果.

4.6.7 定理 若$\bar{x}$是问题(1)的局部最优解，$N_M(\bar{x})\subset\Phi$(Φ依(13))，则$\bar{x}$是式(1)的广义K-T点.

注 在4.1.6的条件下，4.6.7中的条件$N_M(\bar{x})\subset\Phi$与$N_M(\bar{x})=\Phi$相当.

将4.6.5与4.6.7应用到问题(1)的特款

$$\min f(x),\ g(x)\leqslant 0, \tag{14}$$

可得出(亦可直接证明)：

4.6.8 定理 设$Y=\mathbf{R}^m$，$\bar{x}$是问题(14)的局部最优解. 则以下每个条件推出$\bar{x}$是(14)的广义K-T点：(i) $\Phi\overset{\text{def}}{=}\sum_{i\in I}\mathbf{R}_+\partial g_i(\bar{x})$弱*

闭，$\{z : g_i^\circ(\overline{x},z) \leqslant 0(\forall i \in I)\} \subset K_G(\overline{x})$；(ii) $N_G(\overline{x}) \subset \Phi$.

注 若$\forall u_i \in \partial g_i(\overline{x})$：$\{u_i : i \in I\}$线性无关，则由4.1.8知4.6.8中条件(i)，(ii)皆满足，因而(14)的局部最优解$\overline{x}$是广义K-T点. 不过，这一结论已包含在4.6.4中.

§4.4中的结果亦有相应的非光滑推广. 例如，4.4.2可推广为

4.6.9 定理 设Y_+^*有弱*紧凸基；$h(x) = Bx - b$，(f,g)在集$H = h^{-1}(0)$上为$\mathbf{R}_+ \times Y_+$-次类凸，或为$*$拟凸且$*$lsc；存在$\hat{x} \in H$：$\langle Y_+^* \backslash \{0\}, g(\hat{x})\rangle < 0$. 若$\overline{x}$是问题(1)(对于$D = X$)的最优解，则$\overline{x}$是(1)的广义K-T点.

证 在定理条件下，如同4.4.2之证，可应用3.6.6得出$\overline{\lambda} \in Q$，使得

$$f(\overline{x}) = (f + \overline{\lambda}g)(\overline{x}) = \min(f + \overline{\lambda}g)(H).$$

由4.6.6，有$0 \in \partial(f + \overline{\lambda}g)(\overline{x}) + N_H(\overline{x})$. 因$H = N(B) + \overline{x}$，故$T_H(\overline{x}) = K_H(\overline{x}) = N(B)$(参看2.5.2(iv))，从而有$N_H(\overline{x}) = -N(B)^\perp = R(B^*)$，于是

$$\begin{aligned} 0 &\in \partial(f + \overline{\lambda}g)(\overline{x}) + R(B^*) \\ &\subset \partial f(\overline{x}) + \partial g_{\overline{\lambda}}(\overline{x}) + R(B^*), \end{aligned}$$

这表明$\overline{x}$是广义K-T点. □

下面转向考虑充分条件. 首先，利用"广义(F,r)-凸性"(2.2.8)，可建立对应于4.5.1的如下"John充分条件"：

4.6.10 定理 设$(\overline{\rho},\overline{\lambda},\overline{\mu}) \in \mathbf{R}_+ \times Q \times Z^*$满足条件(5)，则以下每个条件推出$\overline{x}$为问题(1)的严格最优解：

(i) $L(\cdot,\overline{\rho},\overline{\lambda},\overline{\mu})$在$\overline{x}$关于$M$为广义$(F,r)$-严格伪凸，$r \geqslant 0$；

(ii) $\overline{\rho}f$与$\overline{\lambda}g + \overline{\mu}h$之一在$x$关于$M$为广义$(F,r)$-严格伪凸，另一个为广义$(F,r')$-拟凸，$r + r' \geqslant 0$；

(iii) $\overline{\rho}f,\overline{\lambda}g,\overline{\mu}h$之一在$\overline{x}$关于$M$为广义$(F,r)$-严格伪凸，另两个分别为广义$(F,r')$-拟凸与广义$(F,r'')$-拟凸，$r + r' + r'' \geqslant 0$.

上述的 F 是给定的且依 1.5.5.

证 不妨设 $\bar{\rho}f$ 在 $\bar{x}$ 关于 M 为广义(F,r)-严格伪凸，$\bar{\lambda}g$ 在 $\bar{\mu}h$ 分别为广义(F,r')-拟凸与广义(F,r'')-拟凸，$r+r'+r''\geqslant 0$，其余情况类似可证. 若 $\bar{x}$ 非严格最优解，则 $\exists x\in M\backslash\{\bar{x}\}: f(x)\leqslant f(\bar{x})$，从而 $\bar{\rho}f(x)\leqslant \rho f(\bar{x})$，于是，由 $\bar{\rho}f$ 在 $\bar{x}$ 广义(F,r)-严格伪凸有

$$\forall u\in \partial f_{\bar{\rho}}(\bar{x}): F(x,u)+r|x-\bar{x}|<0. \tag{15}$$

其次，由 $\bar{\lambda}g(x)\leqslant 0=\bar{\lambda}g(\bar{x})$ 及 $\bar{\lambda}g$ 为广义(F,r')-拟凸有

$$\forall v\in \partial g_{\bar{\lambda}}(\bar{x}): F(x,v)+r'|x-\bar{x}|\leqslant 0. \tag{16}$$

同理

$$\forall w\in \partial h_{\bar{\mu}}(\bar{x}): F(x,w)+r''|x-\bar{x}|\leqslant 0. \tag{17}$$

由(5)及§2.3(5)，有 $u\in \partial f_{\bar{\rho}}(\bar{x}), v\in \partial g_{\bar{\lambda}}(\bar{x}), w\in \partial h_{\bar{\mu}}(\bar{x})$，使 $u+v+w=0$. 这结合(15)～(17)得

$$\begin{aligned} 0 &= F(x,u+v+w) \\ &\leqslant F(x,u)+F(x,v)+F(x,w) \\ &< -(r+r'+r'')|x-\bar{x}|\leqslant 0, \end{aligned}$$

得出矛盾. □

类似地，可建立与 4.5.3 对应的"K-T 充分条件"：

4.6.11 定理 设 $(\bar{\lambda},\bar{\mu})\in Q\times Z^*$ 满足 $0\in\partial_x L(\bar{x},\bar{\lambda},\bar{\mu})$，则以下每个条件推出 $\bar{x}$ 为问题(1)的最优解：

(i) $L(\cdot,\bar{\lambda},\bar{\mu})$ 在 $\bar{x}$ 处关于 M 为广义(F,r)-伪凸，$r\geqslant 0$；

(ii) f 与 $\bar{\lambda}g+\bar{\mu}h$ 在 $\bar{x}$ 处关于 M 分别为广义(F,r)-伪凸与广义(F,r')-拟凸，$r+r'\geqslant 0$；

(iii) $f,\bar{\lambda}g$ 与 $\bar{\mu}h$ 在 $\bar{x}$ 处关于 M 分别为广义(F,r)-伪凸、广义(F,r')-拟凸与广义(F,r'')-拟凸，$r+r'+r''\geqslant 0$.

参考文献：[35,39,89,93,96,97,106,108,116,146,152,179,230].

第五章　对偶理论

写出一个最优化问题

$$(\mathrm{P}):\quad \min f(x)=\alpha,\ x\in M,$$

总意味着 α 是**最优值**，即 $\alpha=\inf f(M)$（这与(P)是否有解无关）. 对偶理论的目标是：给出某种系统的方法，用以对每个给定的**原问题**(P)，能构成一个与之关联的**对偶问题**

$$(\mathrm{P}^*):\quad \max F(y)=\beta,\ y\in N$$

（其中 $\beta=\sup F(N)$），使得在一定条件下可以证明 $\beta\leqslant\alpha$（"弱对偶性"，这通常不成问题），或 $\alpha=\beta$（"强对偶性"），或更强的对偶结论. 这类结论使得对问题(P)与(P^*)的研究能够互相转化，从而导向求解的途径. 例如，若已知(P)与(P^*)分别有解 $\bar{x}$ 与 $\bar{y}$ 且 $\alpha=\beta$，则从方程 $f(\bar{x})=F(\bar{y})$ 可解出 $\bar{x}$（若已知 $\bar{y}$）或 $\bar{y}$（若已知 $\bar{x}$）.

有多种构成对偶问题的方法. 一种常用的方法是，取一适当的函数 $L(x,y)$，使

$$f(x)=\sup L(x,N),\quad F(y)=\inf L(M,y).$$

如上的 $L(x,y)$ 称为"抽象 Lagrange 函数". 可通过不同途径获得所需的 L，采用上一章所述的 Lagrange 函数是常见的选择.

如同上一章一样，对于标准的最优化问题

$$\min f(x)=\alpha,g(x)\leqslant 0,h(x)=0,x\in D,$$

总假定 $f:X\to\mathbf{R},g:X\to Y,h:X\to Z,D\subset X$，$Y$ 中的序 $\leqslant$ 由闭凸锥 Y_+ 导入；$M=D\cap G\cap H,G=g^{-1}(-Y_+),H=h^{-1}(0),Q=Y_+^*\cap\{g(\bar{x})\}^{\perp}$.

§ 5.1 鞍　　点

给定 $\varphi: D \times K \to \mathbf{R}$，令

$$f(x) = \sup \varphi(x,K), \quad F(y) = \inf \varphi(D,y); \tag{1}$$

$$\alpha = \inf f(D), \quad \beta = \sup F(K). \tag{2}$$

如 § 3.5 中已提及的，φ 关联着一对极值问题：

$$(\mathrm{P}): \quad \min f(x) = \alpha, \ x \in D;$$

$$(\mathrm{P}^*): \quad \max F(y) = \beta, \ y \in K.$$

称 $\varphi(x,y)$ 为问题(P)的(抽象)Lagrange 函数；称问题(P^*)为问题(P)关于 Lagrange 函数 φ 的对偶问题. 注意问题(P)是问题(P^*)关于 Lagrange 函数 $-\varphi$ 的对偶问题，因此可以说，问题(P)与问题(P^*)互为对偶问题.

问题(P)与(P^*)恒具弱对偶性，即 $\beta \leqslant \alpha$(参考 § 3.5)；强对偶结论 $\alpha = \beta$ 则并非总能成立. 使 $\alpha = \beta$ 成立的一种特殊情况是出现鞍点.

由 § 3.5 式(5)，$(\bar{x},\bar{y})$ 为 φ 的鞍点意味着

$$\varphi(\bar{x},K) \leqslant \varphi(\bar{x},\bar{y}) \leqslant \varphi(D,\bar{y}). \tag{3}$$

式(3)表明 $\varphi(x,y)$ 在点 $(\bar{x},\bar{y})$“沿 x”取最小而“沿 y”取最大，这正是“鞍点”一词给人的直观印象.

5.1.1 命题　以下条件互相等价：

(i) $(\bar{x},\bar{y})$ 是 φ 的鞍点；

(ii) $\bar{x}$ 与 $\bar{y}$ 分别为问题(P)与(P^*)的解且 $\alpha = \beta$；

(iii) $f(\bar{x}) = F(\bar{y}) (= \varphi(\bar{x},\bar{y}))$.

证明是平凡的.

注意，5.1.1 不用到关于 φ、D、K 的任何特殊假定，因而具有高度的一般性. 5.1.1 表明，一旦求出了 φ 的鞍点，就同时求出了问题(P)与(P^*)的解. 不过，鞍点的出现是极不寻常的.

结合 5.1.1 与 2.1.6 及 1.6.3 直接得出：

5.1.2 命题　设 $D \subset X, K \subset Y, \varphi(\cdot, \bar{y})$ 与 $\varphi(\bar{x}, \cdot)$ 分别为凸函数与凹函数，则 $(\bar{x}, \bar{y})$ 是 φ 的鞍点的充要条件是

$$\begin{cases} 0 \in \partial_x \varphi(\bar{x}, \bar{y}) + \partial \delta_D(\bar{x}), \\ 0 \in \partial_y [-\varphi(\bar{x}, \bar{y})] + \partial \delta_K(\bar{y}). \end{cases} \tag{4}$$

若 G- 导数 $\varphi_x(\bar{x}, \bar{y})$ 与 $\varphi_y(\bar{x}, y)$ 存在，则 $(\bar{x}, \bar{y})$ 为 φ 的鞍点 的充要条件是

$$\langle \varphi_x(\bar{x}, \bar{y}), D - \bar{x} \rangle \geqslant 0 \geqslant \langle \varphi_y(\bar{x}, \bar{y}), K - \bar{y} \rangle. \tag{5}$$

当 $\bar{x} \in D^\circ, \bar{y} \in K^\circ$ 时，条件式(5) 相当于

$$\varphi_x(\bar{x}, \bar{y}) = 0, \quad \varphi_y(\bar{x}, \bar{y}) = 0. \tag{6}$$

注意，条件式(5) 与(6) 作为鞍点的必要条件无需关于 φ 的凸性假设. 条件式(6) 可看作某种抽象的"K-T 条件"，在某些具体场合它实际上就是通常的 K-T 条件(参看下节).

下面利用 §3.5 的结果导出鞍点存在的某些充分条件. 如在 §3.5 中已提到的，§3.5 中的等式(3) 与(4) 成立分别意味着问题(P) 有解且 $\alpha = \beta$ 与(P^*) 有解且 $\alpha = \beta$. 因此，§3.5 之式(3)、式(4) 同时成立 $\Leftrightarrow$ 问题(P) 与(P^*) 均有解且 $\alpha = \beta \Leftrightarrow \varphi$ 有鞍点(5.1.1). 于是，组合 3.5.8 ～ 3.5.10 的条件，可得 φ 有鞍点的一系列充分条件，其主要者列举如下(沿用 §3.5 的记号).

5.1.3 定理　以下每个条件推出 φ 有鞍点：

(i) D, K 为 Hausdorff 空间，φ 次类鞍且半连续，存在 $\omega \subset D$, $\omega' \subset K, \sigma < \alpha, \sigma' > \beta$，使 $D(\omega', \sigma')$ 与 $K(\omega, \sigma)$ 为紧集.

(ii) D, K 为紧 Hausdorff 空间，φ 近次类鞍且半连续.

(iii) X, Y 为 LCS，$D \subset X$ 与 $K \subset Y$ 凸紧，φ 拟鞍且半连续.

(iv) X, Y 为自反 B- 空间，$D \subset X$ 与 $K \subset Y$ 弱闭，φ 次类鞍且对 x 为 wlsc 而对 y 为 wusc，$\exists \lambda \in F_+^+, \rho \in E_+^+$，使 $\sum\limits_y \lambda(y) = \sum\limits_x \rho(x) = 1$，且

$$\liminf_{x \in D, |x| \to \infty} \langle \lambda, \varphi(x, \cdot) \rangle > \beta, \tag{7}$$

$$\limsup_{y \in K, |y| \to \infty} \langle \rho, \varphi(\cdot, y) \rangle < \alpha. \tag{8}$$

可用一些较强但更简单的条件来替换 5.1.3 中的条件.例如，"φ 近次类鞍" 可换成"φ 类鞍" 甚至"φ 鞍";"D 与 K 弱闭" 可换成"D 与 K 凸闭"; 条件式(7)与式(8)可分别换成：

$$\exists \lambda \in F_+^*: \lim_{x\in D,\ |x|\to\infty} \langle \lambda, \varphi(x,\cdot)\rangle = \infty;$$

$$\exists \rho \in E_+^*: \lim_{y\in K,\ |y|\to\infty} \langle \rho, \varphi(\cdot,y)\rangle = -\infty.$$

后者又可以换成更简单的条件：

$$\exists \hat{y} \in K: \lim_{x\in D,\ |x|\to\infty} \varphi(x,\hat{y}) = \infty; \tag{9}$$

$$\exists \hat{x} \in D: \lim_{y\in K,\ |y|\to\infty} \varphi(\hat{x},y) = -\infty. \tag{10}$$

于是由 5.1.3 得出：

5.1.4 推论 设 X,Y 为自反 B- 空间,$D\subset X$ 与 $K\subset Y$ 为凸闭集,φ 为半连续的鞍函数,且条件式(9) 与式(10) 满足,则 φ 有鞍点.

这一结果在[238] 中(参看[238,Th.49A]) 是从完全不同的另一途径得到的,它不过是 5.1.3 的一种很特殊的情况.

5.1.5 定义 若$(\bar{x},\bar{y}) \in D\times K$ 与 $\varepsilon \geqslant 0$ 满足

$$\varphi(\bar{x},K) \leqslant \varphi(D,\bar{y}) + \varepsilon, \tag{11}$$

则称 $(\bar{x},\bar{y})$ 为 φ 的 ε- 鞍点.

直接从式(11)推出

$$\alpha \leqslant \sup\varphi(\bar{x},K) \leqslant \inf\varphi(D,\bar{y}) + \varepsilon \leqslant \beta + \varepsilon;$$

$$\varphi(\bar{x},\bar{y}) - \varepsilon \leqslant \alpha \leqslant \beta + \varepsilon \leqslant \varphi(\bar{x},\bar{y}) + 2\varepsilon,$$

可见 $\alpha,\beta \neq \pm\infty$. 因此,$\alpha = \beta \neq \pm\infty \Leftrightarrow \forall \varepsilon > 0$,$\varphi$ 有 ε- 鞍点. 注意"0- 鞍点" 就是鞍点.

关于 ε- 鞍点有类似于 5.1.3 的以下结果.

5.1.6 定理 设 α 有限,$\varepsilon > 0$. 假定:(i) φ 对 x 几乎类凸而对 y 次类凹;(ii) 存在 $r \in (\alpha-\varepsilon,\alpha)$ 与 F 的 0 邻域 V,使 $N_r \cap V$ 闭 (N_r 依 §3.5(9)),且问题(P) 有解,则对(P) 的任何解 $\bar{x}$,有 $\bar{y} \in K$,使$(\bar{x},\bar{y})$ 为 φ 的 ε- 鞍点. 条件(ii) 可代以如下三条件之一:(iii) $\exists\delta > 0,\forall r \in (\alpha-\delta,\alpha]$,存在 F 的 0- 邻域 V,使 $N_r \cap V$ 闭;(iv) D 为 Hausdorff 空间,φ 对 x 为 lsc;$\exists\,\omega \subset K,\sigma > \alpha$,使 $D(\omega,\sigma)$ 为

紧集;(v) X 为自反 B- 空间,$D\subset X$ 弱闭,φ 对 x 为 wlsc;$\exists\lambda\in F_+^*$,使 $\sum_y\lambda(y)=1$,且

$$\liminf_{x\in D,\,|x|\to\infty}\langle\lambda,\varphi(x\,\cdot)\rangle>\alpha. \tag{12}$$

证 因 $r+\varepsilon-\alpha>0$,故由 3.5.6 有 $\bar{y}\in K$,使得

$$r\leqslant\varphi(D,\bar{y})+r+\varepsilon-\alpha.$$

若 $\bar{x}$ 是(P) 的解,则 $\alpha=\sup\varphi(\bar{x},K)$. 以此代入上面的不等式得出式(11),即$(\bar{x},\bar{y})$ 是 φ 的 ε- 鞍点.

$\forall n\geqslant 1$,取 $x_n\in D$,使 $f(x_n)<\alpha+n^{-1}$,即 $\varphi(x_n,\cdot)<\alpha+n^{-1}$. 设条件(iii) 满足,则 n 充分大时

$$\frac{1}{n}=[\varphi(x_n,\cdot)-\alpha]+[\alpha+\frac{1}{n}-\varphi(x_n,\cdot)]\in N_\alpha\cap V.$$

这推出 $0\in N_\alpha$,即有 $\bar{x}\in D$,使 $\varphi(\bar{x},K)\leqslant\alpha$,从而 $\alpha=\sup\varphi(\bar{x},K)$,即 $\bar{x}$ 是问题(P) 的解. 可见(iii)→(ii).

由 3.5.3 知(iv)⇒(iii),(v)⇒(iii). 于是定理得证. □

如同在 5.1.3 中一样,条件(12) 可代以更简单的条件(9).

与 5.1.6 对偶的结果可直接写出如下.

5.1.7 定理 设 β 有限,$\varepsilon>0$. 假定:(i) φ 对 x 次类凸而对 y 几乎类凹;(ii) 存在 $r\in(\beta,\beta+\varepsilon)$ 与 E 的 0- 邻域 U,使 $M_r\cap U$ 闭(M_r 依 §3.5(9)),且问题(P^*) 有解. 则对(P^*) 的任何解 $\bar{y}$,有 $\bar{x}\in D$,使$(\bar{x},\bar{y})$ 为 φ 的 ε- 鞍点. 条件(ii) 可代以如下条件之一:(iii) $\exists\delta>0,\forall r\in[\beta,\beta+\delta)$,存在 E 的 0- 邻域 U,使 $M_r\cap U$ 为闭集;(iv) K 为 Hausdorff 空间,φ 对 y 为 usc;$\exists\omega\subset D,\sigma<\beta$;$K(\omega,\sigma)$ 紧;(v) Y 为自反 B- 空间,$K\subset Y$ 弱闭,φ 对 y 为 wusc,存在 $\rho\in E_+^*$,使 $\sum_x\rho(x)=1$,且

$$\limsup_{y\in K,\,|y|\to\infty}\langle\rho,\varphi(\cdot,y)\rangle<\beta. \tag{13}$$

其中条件式(13)可代以条件式(10).

参考文献:[93,104,113,192,238].

§ 5.2 Lagrange 对偶

今将上节结论用于问题

$$\min f(x)=\alpha,\quad g(x)\leqslant 0,\quad h(x)=0,\quad x\in D. \tag{1}$$

Lagrange 函数 $L(x,\lambda,\mu)=f(x)+\lambda g(x)+\mu h(x)$ 将起 § 5.1 中的 $\varphi(x,y)$ 的作用.

令 $M=D\cap G\cap H$. 首先指出如下结论:

1 利用1.4.2 (i) 易得 $\sup L(x,Y_+^*,Z^*)=f(x)+\delta_M(x)$, 因此

$$\inf_{x\in D}\sup L(x,Y_+^*,Z^*)=\inf_{x\in M}f(x)=\alpha.$$

可见,借助于 $L(x,\lambda,\mu)$ 可将问题(1)的对偶问题表为

$$\max\inf L(D,\lambda,\mu)=\beta,(\lambda,\mu)\in Y_+^*\times Z^*. \tag{1*}$$

称问题(1*)为(1)的 **Lagrange 对偶**. 注意(1*)的约束比较简单, 目标函数 $\inf L(D,\lambda,\mu)$ 则未必如此.

2 极值关系: $f(\bar{x})+\delta_M(\bar{x})=\inf L(D,\bar{\lambda},\bar{\mu})$, 这相当于

$$\begin{cases}f(\bar{x})=\min L(D,\bar{\lambda},\bar{\mu}),\\ \langle\bar{\lambda},g(\bar{x})\rangle=0,(\bar{x},\bar{\lambda},\bar{\mu})\in M\times Y_+^*\times Z^*.\end{cases} \tag{2}$$

结合以上结论与 5.1.1 得出:

5.2.1 定理 设 $(\bar{x},\bar{\lambda},\bar{\mu})\in M\times Y_+^*\times Z^*$.

(i) $(\bar{x},\bar{\lambda},\bar{\mu})$ 是 L 的鞍点 $\Leftrightarrow \bar{x}$ 与 $(\bar{\lambda},\bar{\mu})$ 分别为问题(1)与(1*)的解且 $\alpha=\beta\Leftrightarrow$ 关系式(2)成立.

(ii) 若 $(\bar{x},\bar{\lambda},\bar{\mu})$ 是 L 的鞍点, D 凸, f,g,h 在 $\bar{x}$ 可微, 则

$$\langle L_x(\bar{x},\bar{\lambda},\bar{\mu}),D-\bar{x}\rangle\geqslant 0=\langle\bar{\lambda},g(\bar{x})\rangle. \tag{3}$$

因此当 $\bar{x}\in D$ 时 $\bar{x}$ 是 K-T 点.

(iii) 若 $\bar{x}$ 是问题(1)的 K-T 点, $\bar{\lambda},\bar{\mu}$ 是 K-T 乘子, $L(\cdot,\bar{\lambda},\bar{\mu})$ 在 $\bar{x}$ 关于 D 伪凸, 则 $(\bar{x},\bar{\lambda},\bar{\mu})$ 是 L 的鞍点.

证 (i) 直接从 5.1.1 得出. 利用 1.6.3 从式(2) 推出式(3). 在 (iii) 的条件下, 式(3) 必成立; 利用伪凸假设及 1.6.3, 从式(3) 推出式(2), 故$(\bar{x},\bar{\lambda},\bar{\mu})$ 是鞍点. □

5.2.1 可与 §4.3 的结果对照. 在 5.2.1 中, 鞍点条件似乎起了"约束品性"的作用.

5.2.2 定理 设 D 凸, $(\bar{x},\bar{\lambda},\bar{\mu})\in M\times Y_+^*\times Z^*$, f,g,h 在 $\bar{x}$ 邻近为 Lip.

(i) 若$(\bar{x},\bar{\lambda},\bar{\mu})$ 是 Lagrange 函数 L 的鞍点, 则

$$0\in\partial_x L(\bar{x},\bar{\lambda},\bar{\mu})+\partial\delta_D(\bar{x}),\langle\bar{\lambda},g(\bar{x})\rangle=0. \tag{4}$$

(ii) 若条件(4)满足, 且函数 $L(\cdot,\bar{\lambda},\bar{\mu})$ 在 $\bar{x}$ 关于 D 为广义伪凸, 则$(\bar{x},\bar{\lambda},\bar{\mu})$ 是 L 的鞍点.

5.2.2 可看作是 5.2.1 的非光滑类似.

证 由 2.2.9, 式(2)蕴涵式(4); 在广义伪凸的假设下, 式(4)蕴涵式(2), 由此得所要证之结果. □

对于问题(1)的特款

$$\min f(x)=\alpha,\quad g(x)\leqslant 0,\ x\in D, \tag{5}$$

对偶问题(1*)简化成

$$\max\inf L(D,\lambda)=\beta,\ \lambda\in Y_+^*, \tag{5*}$$

其中 $L(\cdot,\lambda)=f+\lambda g$. 令 $M=D\cap G$, 直接从 5.2.1, 5.2.2 得出:

5.2.3 推论 设$(\bar{x},\bar{\lambda})\in M\times Y_+^*$, 则有以下结论:

(i) $(\bar{x},\bar{\lambda})$ 是 L 的鞍点 $\Leftrightarrow$ $\bar{x}$ 与 $\bar{\lambda}$ 分别为问题(5) 与(5*) 的解且 $\alpha=\beta\Leftrightarrow$ 以下关系成立:

$$f(\bar{x})=\min L(D,\bar{\lambda}),\quad\langle\bar{\lambda},g(\bar{x})\rangle=0. \tag{6}$$

(ii) 若 $(\bar{x},\bar{\lambda})$ 是 L 的鞍点, D 凸, f,g 在 $\bar{x}$ 可微, 则

$$\langle L_x(\bar{x},\bar{\lambda}),D-\bar{x}\rangle\geqslant 0=\langle\bar{\lambda},g(\bar{x})\rangle. \tag{7}$$

因此当 $\bar{x}\in \mathring{D}$ 时 $\bar{x}$ 是K-T 点. 反之, 若 $\bar{x}$ 是K-T 点, $\bar{\lambda}$ 是K-T 乘子, $L(\cdot,\bar{\lambda})$ 在 $\bar{x}$ 关于 D 为伪凸, 则$(\bar{x},\bar{\lambda})$ 是 L 的鞍点.

(iii) 若$(\bar{x},\bar{\lambda})$ 是 L 的鞍点, D 凸, f,g 在 $\bar{x}$ 邻近为 Lip, 则

$$0\in\partial_x L(\bar{x},\bar{\lambda})+\partial\delta_D(\bar{x}),\quad\langle\bar{\lambda},g(\bar{x})\rangle=0. \tag{8}$$

反之，若式(8)满足且 $L(\cdot,\bar{\lambda})$ 在 $\bar{x}$ 关于 D 为广义伪凸，则$(\bar{x},\bar{\lambda})$ 是 L 的鞍点.

以 $L(x,\lambda)$ 代 $\varphi(x,y)$，从 3.5.7 ~ 3.5.9，5.1.3，5.1.6，5.1.7 得出：

5.2.4 定理 设(f,g) 在集 D 上为 $\mathbf{R}_+\times Y_+$ - 类凸. 给定条件：

(i) $f|D$ 与 $g|D$ 分别为 lsc 与 $*$lsc；

(ii) $f|D$ 与 $g|D$ 分别为 wlsc 与 $*$wlsc；

(iii) 存在有限集 $\omega\subset D$ 与 $\sigma\in\mathbf{R}$，使

$$\{\lambda\in Y_+^*:L(\omega,\lambda)\geqslant\sigma\}$$

为弱*紧集；

(iv) 存在有限集 $\Lambda\subset Y_+^*$ 与 $\sigma\in\mathbf{R}$，使

$$\{x\in D:L(x,\Lambda)\leqslant\sigma\}$$

为弱紧集；

(v) X 为自反 B-空间，D 弱闭，且

$$\exists\hat{\lambda}\in Y_+^*:\liminf_{x\in D,\|x\|\to\infty}L(x,\hat{\lambda})>\alpha;\tag{9}$$

(vi) Y 为自反 B-空间，且

$$\exists\hat{x}\in D:\limsup_{\lambda\in Y_+^*,\|\lambda\|\to\infty}L(\hat{x},\lambda)<\beta.\tag{10}$$

则有以下结论：

1° 若条件(i)，(iv)(其中 $\sigma>\beta$) 或(ii)，(v) 满足，则问题(5)有解且 $\alpha=\beta\neq\pm\infty$.

2° 若条件(iii)(其中 $\sigma<\alpha$) 或(vi) 满足，则问题(5^*) 有解且 $\alpha=\beta\neq\pm\infty$.

3° 若 D 弱紧且条件(i) 满足，则问题(5) 有解且 $\alpha=\beta>-\infty$；条件"(f,g) 类凸"可代以条件"$L(x,\lambda)$ 对 x 拟凸".

4° 若 α 有限，条件(i)，(iv)(其中 $\sigma>\alpha$) 或(ii)，(v) 满足，则问题(5) 有解 $\bar{x}$，且 $\forall\varepsilon>0$，$\exists\lambda_\varepsilon\in Y_+^*$，使$(\bar{x},\lambda_\varepsilon)$ 是 $L(x,\lambda)$ 在 $D\times Y_+^*$ 上的 ε-鞍点.

5° 若 β 有限，条件(iii)(其中 $\sigma<\beta$) 或(vi) 满足，则问题(5^*)

有解 $\bar\lambda$，且 $\forall \varepsilon > 0, \exists x_\varepsilon \in D$，使 $(x_\varepsilon, \bar\lambda)$ 是 $L(x,\lambda)$ 在 $D \times Y_+^*$ 上的 ε- 鞍点.

6° 若 1° 与 2° 的条件均满足，则 $L(x,\lambda)$ 在 $D \times Y_+^*$ 上有鞍点.

证 只需指明以下事实：(f,g) 为 $\mathbf{R}_+ \times Y_+$- 类凸 $\Rightarrow L(x,\lambda)$ 对 x 类凸；条件(i) $\Rightarrow L(x,\lambda)$ 对 x 为 lsc；条件(ii) $\Rightarrow L(x,\lambda)$ 对 x 为 wlsc. □

注 条件(9)与(10)可分别换成更简单的条件

$$\liminf_{x \in D, |x| \to \infty} f(x) > \alpha$$

与

$$\exists \hat{x} \in D;\ g(\hat{x}) \ll 0.$$

将 5.2.3(iii)用到的"广义伪凸"换成某种更一般的广义凸性，得到如下鞍点定理.

5.2.5 定理 设 $(\bar{x}, \bar\lambda) \in M \times Q, 0 \in \partial_x L(\bar{x}, \lambda), f, g$ 为局部 Lip，且以下条件之一满足：(i) $L(\cdot, \bar\lambda)$ 在 $\bar{x}$ 关于 D 为广义 (F,r)-伪凸，$r \geqslant 0$，(ii) f 与 λg 在 $\bar{x}$ 关于 D 分别为广义 (F,r_1)- 凸与广义 (F,r_2)- 凸，$r_1 + r_2 \geqslant 0$(F 是取定的且如 1.5.5). 则 $(\bar{x}, \bar\lambda)$ 是 $L(x,\lambda)$ 在 $D \times Y_+^*$ 上的鞍点.

证 首先，可直接看出对任何 $\lambda \in Y_+^*$ 有

$$L(\bar{x}, \lambda) = f(\bar{x}) + \lambda g(\bar{x}) \leqslant f(\bar{x}) = L(\bar{x}, \bar\lambda),$$

即 $L(\bar{x}, Y_+^*) \leqslant L(\bar{x}, \bar\lambda)$. 若 $(\bar{x}, \bar\lambda)$ 不是 $L(x,\lambda)$ 的鞍点，则有 $x \in D$ 使 $L(x, \bar\lambda) < L(\bar{x}, \bar\lambda)$. 若条件(i) 满足，则由 $0 \in \partial_x L(\bar{x}, \bar\lambda)$ 有(参看 2.2.8)

$$0 > F(x, 0) + r|x - \bar{x}| = r|x - \bar{x}| \geqslant 0,$$

得出矛盾. 若条件(ii)满足，取 $u \in \partial f(\bar{x}), v \in \partial(\bar\lambda g)(\bar{x})$，使 $0 = u + v \in \partial f(\bar{x}) + \partial(\bar\lambda g)(\bar{x}) \supset \partial_x L(\bar{x}, \bar\lambda)$(§2.3(5))，则

$$\begin{aligned}
0 = F(x, u+v) &\leqslant F(x,u) + F(x,v) \\
&\leqslant -r_1|x - \bar{x}| + f(x) - f(\bar{x}) \\
&\quad - r_2|x - \bar{x}| + \bar\lambda g(x) - \bar\lambda g(\bar{x}) \\
&= -(r_1 + r_2)|x - \bar{x}| + L(x, \bar\lambda) - L(\bar{x}, \bar\lambda) < 0,
\end{aligned}$$

亦得矛盾. 故 $(\bar{x}, \bar\lambda)$ 是 $L(x,\lambda)$ 的鞍点. □

5.2.5 可看作 [108, Th. 3.1] 的某种推广.

导出对偶定理的另一途径是应用择一定理，其思路颇类似于 §4.4 中的考虑. 设 α 有限，令 $\tilde{f}(x)=f(x)-\alpha$. 则 α 是问题(1)与问题(5)的最优值就分别意味着问题

$$\exists x\in D: \tilde{f}(x)<0, g(x)\leqslant 0, h(x)=0 \tag{11}$$

与

$$\exists x\in D: \tilde{f}(x)<0, g(x)\leqslant 0 \tag{12}$$

无解. 这就自然联系到备择的问题：

$$\exists(\lambda,\mu)\in Y_+^*\times Z^*: (\tilde{f}+\lambda g+\mu h)(D)\geqslant 0 \tag{13}$$

与

$$\exists\lambda\in Y_+^*: (\tilde{f}+\lambda g)(D)\geqslant 0. \tag{14}$$

若问题(11)与问题(13)两择一，则问题(13)必有解 $(\lambda,\mu)=(\bar{\lambda},\bar{\mu})$，这意味着 $\alpha\leqslant L(D,\bar{\lambda},\bar{\mu})$，从而 $\alpha\leqslant\inf L(D,\bar{\lambda},\bar{\mu})\leqslant\beta$，于是 $\alpha=\beta$ 且 $(\bar{\lambda},\bar{\mu})$ 是问题(1^*)的解. 基于以上分析，从 3.4.2 推出：

5.2.6 定理　设 α 有限，f,g,h 满足 3.4.2 的条件(i)～(iv)(取 $W=\mathbf{R}$)，则问题(1^*)有解且 $\alpha=\beta$.

类似地，利用 3.4.8 与 3.6.6 获得问题(12)与问题(14)两择一的条件，然后得出：

5.2.7 定理　设 α 有限；$\exists\bar{x}\in D: \langle Y_+^*\setminus\{0\}, g(\bar{x})\rangle<0$. 假定以下两条件之一满足：(i) $Y_+^{*\circ}\neq\varnothing$，$(f,g)$ 在 D 上为 $\mathbf{R}_+\times Y_+$-次类凸；(ii) Y_+^* 有弱*紧凸基，(f,g) 在 D 上为 $\mathbf{R}_+\times Y_+$-次类凸，或为 * 拟凸且 * lsc，则问题(5^*)有解且 $\alpha=\beta$.

5.2.7 综合并推广了[85, 104, 114]中的一些结果.

在 5.2.7 的条件下，若再假定 $\bar{x}$ 是问题(5)的最优解，f,g 在 $\bar{x}$ 可微，D 凸且 $\bar{x}\in D^\circ$，则由 5.2.3 得出 $\bar{x}$ 是问题(5)的 K-T 点. 这一结论补充了 4.3.6. 显然，5.2.6 亦有类似的推论.

在 5.2.6(部分地 5.2.7)中，$Y_+^{*\circ}\neq\varnothing$ 是重要的. 倘无此条件，则可用 3.4.12 得如下结果：

5.2.8 定理　设 $\varphi\overset{\text{def}}{=\!=}(f,g,h)$ 在 D 上为 $\mathbf{R}_+\times Y_+\times 0$-次类凸，问题(1)的最优值 α 有限，且以下条件之一满足：(i) 存在 $\sigma>\alpha$ 与 $Y\times Z$ 的 0-邻域 V，使 $C\cap((-\infty,\sigma]\times V)$ 闭，$C=\varphi(D)+$

$\mathbf{R}_+ \times Y_+ \times 0$;(ii) $\{(x,r,y,h(x)): x\in D, f(x)\leqslant \alpha+r, g(x)\leqslant y\}$ 为闭集,存在 $\sigma>\alpha$ 与空间 Y,Z 的闭 0- 邻域 U,V,使得 $D\cap f^{-1}((-\infty,\sigma])\cap g^{-1}(U-Y_+)\cap h^{-1}(V)$ 为紧集;(iii) X 为自反 B-空间,D 弱闭,f 为 wlsc,g 为 * wlsc,在 D 中 $x_n\rightharpoonup x \Rightarrow h(x_n)\rightharpoonup h(x)$,且有 $(\rho,\lambda,\mu)\in \mathbf{R}_+\times Y_+^*\times Z^*$,使得

$$\liminf_{x\in D,\|x\|\to\infty} L(x,\rho,\lambda,\mu) > \rho\alpha.$$

则问题(1)有最优解 $\bar{x}$;$\forall\varepsilon>0,\exists(\lambda_\varepsilon,\mu_\varepsilon)\in Y_+^*\times Z^*$:$(\bar{x},\lambda_\varepsilon,\mu_\varepsilon)$ 是 $L(x,\lambda,\mu)$ 的 ε- 鞍点(5.1.5),因而 $\alpha=\beta$.

证 可指明(ii)⇒(i),(iii)⇒(i)(参考 3.4.9 之证),因此不妨设条件(i)满足.取 $x_n\in M$,使 $f(x_n)<\alpha+n^{-1}(\forall n\geqslant 1)$.不妨设 $\alpha+n^{-1}\leqslant\sigma$,于是 $(\alpha+n^{-1},0,0)\in C\cap((-\infty,\sigma]\times V)$,从而由条件(i)推出 $(\alpha,0,0)\in C$.这意味着 $\exists\bar{x}\in D$:$f(\bar{x})\leqslant\alpha, g(\bar{x})\leqslant 0, h(\bar{x})=0$,可见 $\bar{x}$ 是问题(1)的最优解.

因问题(11)无解,由 3.4.12,$\forall\varepsilon>0,\exists(\lambda_\varepsilon,\mu_\varepsilon)\in Y_+^*\times Z^*$,使得 $L(D,\lambda_\varepsilon,\mu_\varepsilon)>\alpha-\varepsilon$.这推出

$$\sup L(\bar{x},Y_+^*,Z^*) = f(\bar{x}) = \alpha \leqslant \inf L(D,\lambda_\varepsilon,\mu_\varepsilon)+\varepsilon,$$

可见 $(\bar{x},\lambda_\varepsilon,\mu_\varepsilon)$ 是 $L(x,\lambda,\mu)$ 的 ε- 鞍点. □

取 $Z=\{0\}$,从 5.2.8 得出推论:

5.2.9 推论 设 $\varphi\stackrel{\text{def}}{=}(f,g)$ 在 D 上为 $\mathbf{R}_+\times Y_+$-次类凸,问题(5)的最优值 α 有限,且以下条件之一满足:(i) 存在 $\sigma>\alpha$ 与 Y 的 0- 邻域 V,使 $(\varphi(D)+\mathbf{R}_+\times Y_+)\cap((-\infty,\sigma]\times V)$ 闭;(ii) $\{(x,r,y): x\in D, f(x)\leqslant\alpha+r, g(x)\leqslant y\}$ 为闭集,存在 $\sigma>\alpha$ 与 Y 的闭 0- 邻域 V,使 $D\cap f^{-1}((-\infty,\sigma])\cap g^{-1}(V-Y_+)$ 为紧集;(iii) X 为自反 B- 空间,D 弱闭,f 为 wlsc,g 为 * wlsc,且有 $(\rho,\lambda)\in\mathbf{R}_+\times Y_+^*$,使得

$$\liminf_{x\in D,|x|\to\infty} \rho f(x)+\lambda g(x) > \rho\alpha,$$

则问题(5)有最优解 $\bar{x}$;$\forall\varepsilon>0,\exists\lambda_\varepsilon\in Y_+^*$,使 $(\bar{x},\lambda_\varepsilon)$ 为 $L(x,\lambda)$ 的 ε-鞍点,因而 $\alpha=\beta$.

参考文献:[85,93,104,108,114,195,196,238].

§5.3 共轭泛函

为描述一些新的对偶，需要共轭泛函这一工具. 以下设 $f: X \to \overline{\mathbf{R}}, D = D_f \neq \varnothing$.

5.3.1 定义 对任给 $u \in X^*$，令

$$f^*(u) = \sup_{x \in X} [u(x) - f(x)]. \tag{1}$$

如此得一泛函 $f^*: X^* \to \overline{\mathbf{R}}$，称为 f 的**共轭泛函**. 称 $f^{**} \stackrel{\text{def}}{=} (f^*)^*$ 为 f 的二次共轭；称 $\bar{f} \stackrel{\text{def}}{=} f^{**}|X$ 为 f 闭包.

由式(1) 推出 $u(x) - f(x) \leqslant f^*(u)$，即

$$u(x) \leqslant f(x) + f^*(u), (x, u) \in X \times X^*. \tag{2}$$

式(2)称为 **Young 不等式**. 容易验证，仅当 $u \in \partial f(x)$ 时式(2) 成为等式. 特别，若 f 凸且 $\mathrm{D}f(x)$ 存在，则

$$\langle \mathrm{D}f(x), x \rangle = f(x) + f^*(\mathrm{D}f(x)). \tag{3}$$

5.3.2 例 1° 设 $p > 1, p^{-1} + q^{-1} = 1, f(x) = p^{-1}|x|^p (x \in X)$，则对任给 $u \in X^*$ 有

$$\begin{aligned} f^*(u) &= \sup_{x \in X} [u(x) - p^{-1}|x|^p] \\ &\leqslant \sup_{x \in X} (|u||x| - p^{-1}|x|^p) \\ &= \sup_{0 \leqslant t < \infty} (|u|t - p^{-1}t^p) = q^{-1}|u|^q. \end{aligned}$$

另一方面

$$\begin{aligned} f^*(u) &\geqslant \sup_{|x|^{p-1} = |u|} \left[|x| u\left(\frac{x}{|x|}\right) - \frac{1}{p}|x|^p \right] \\ &= |u|^{1/(p-1)}|u| - \frac{1}{p}|u|^{p/(p-1)} = \frac{1}{q}|u|^q, \end{aligned}$$

因此 $f^*(u) = q^{-1}|u|^q$. 以此代入式(2) 得

$$u(x) \leqslant \frac{1}{p}|x|^p + \frac{1}{q}|u|^q, (x, u) \in X \times X^*.$$

特别，取 $x,u\in\mathbf{R}$ 得经典意义下的 Young 不等式：

$$ux\leqslant\frac{1}{p}|x|^p+\frac{1}{q}|u|^q.$$

2° 设 $f(x)=|x|(x\in X)$. 任给 $u\in X^*$，若 $|u|\leqslant 1$，则

$$f^*(u)=\sup_{x\in X}[u(x)-|x|]=0.$$

若 $|u|>1$，则

$$f^*(u)\geqslant\sup_{x\in X}u(x)=\sup_{x\in X}|x|u\left(\frac{x}{|x|}\right)=\infty.$$

故得 $f^*=\delta_B$（参考 2.1.2 之 4°），$B\subset X^*$ 是单位闭球.

3° 设 X 自反，$A=A^*\in L(X,X^*)$ 强单调（1.3.2），$f(x)=\frac{1}{2}\langle Ax,x\rangle$. 则 $f'(x)=Ax$，因此 $f(x)$ 严格凸（1.5.3(ii)）. 任给 $u\in X^*$，$f(x)-u(x)$ 有唯一极小点 x，它满足 $Ax-u=0$（1.6.3），从而 $x=A^{-1}u$（由 1.3.7，A^{-1} 存在且亦为强单调线性算子）. 于是

$$\begin{aligned}f^*(u)&=\sup_{x\in X}[u(x)-f(x)]\\&=-\min_{x\in X}[f(x)-u(x)]\\&=-f(A^{-1}u)+\langle u,A^{-1}u\rangle,\end{aligned}$$

故得 $f^*(u)=\frac{1}{2}\langle u,A^{-1}u\rangle$.

4° 若 $f\in X^*$，则 $f^*=\delta_{\{f\}}$.

5° 对任给 $A\subset X$，有 $\delta_A^*=s_A$，$s_A(u)=\sup u(A)$（参照2.2.2）. 若 A 是锥，则 $\delta_A^*=\delta_{-A^*}$.

"共轭运算"$f\to f^*$ 的以下性质是明显的：

$$(f+v+c)^*(u)=f^*(u-v)-c\ (u,v\in X^*,c\in\mathbf{R});\quad(4)$$

$$(\rho f(ax+b))^*(u)=\rho f^*\left(\frac{u}{a\rho}\right)-\frac{1}{a}u(b)\ (\rho>0,a\neq 0).\quad(5)$$

由式(5)特别得出：

$$(\rho f)^*(u)=\rho f^*\left(\frac{u}{\rho}\right),\quad(6)$$

$$(f(ax+b))^*(u) = f^*\left(\frac{u}{a}\right) - \frac{1}{a}u(b)\ (p > 0, a \neq 0).$$

5.3.3 定理 (i) 若 f 凸且 lsc，则 $f^* \not\equiv \infty$；$f = \bar{f} \Leftrightarrow f$ 凸且 lsc.

(ii) 若 $\bar{f} \leqslant \varphi \leqslant f$，$\varphi$ 凸且 lsc，则 $\bar{f} = \varphi$.

(iii) $f^* = (\bar{f})^*$；当 $f^* \not\equiv \infty$ 时 $\operatorname{epi} \bar{f} = \overline{\operatorname{co}} \operatorname{epi} f$.

证 (i) 首先设 f 凸且 lsc. 取 $\bar{x}, \bar{t}$ 使 $\bar{t} < f(\bar{x}) < \infty$. 因 $\operatorname{epi} f$ 闭凸且 $(\bar{x}, \bar{t}) \bar{\in} \operatorname{epi} f$，由 1.1.2 有 $(u, \alpha) \in X^* \times \mathbf{R}$，使得 $f(x) \leqslant t \Rightarrow u(\bar{x}) + \alpha\bar{t} < u(x) + \alpha t$. 这推出 $\alpha > 0$，$f^*(-u/\alpha) < \infty$，可见 $f^* \not\equiv \infty$. 直接看出 $\bar{f} \leqslant f$. 若有 x_0 使 $t_0 \overset{\text{def}}{=} \bar{f}(x_0) < f(x_0)$，则再用 1.1.2 得 $v \in X^*$，$\beta, r \in \mathbf{R}$，使得

$$f(x) \leqslant t \Rightarrow v(x_0) + \beta t_0 < r < v(x) + \beta t. \tag{7}$$

必定 $\beta \geqslant 0$. 若 $\beta > 0$，则由式(7)推出

$$\bar{f}(x_0) \geqslant \langle -v/\beta, x_0 \rangle - f^*(-v/\beta) > t_0,$$

得出矛盾. 若 $\beta = 0$，则由(7)有 $v(x_0) < r < v(D)$. 取 $u \in X^*$ 使 $f^*(u) < \infty$. 则当 $t \uparrow \infty$ 时有

$$\begin{aligned}
\bar{f}(x_0) &\geqslant \langle u - tv, x_0 \rangle - f^*(u - tv) \\
&= \langle u - tv, x_0 \rangle - \sup_{x \in D} [u(x) - tv(x) - f(x)] \\
&\geqslant \langle u - tv, x_0 \rangle - \sup_{x \in D} [u(x) - tr - f(x)] \\
&= u(x_0) + t[r - v(x_0)] - f^*(u) \to \infty,
\end{aligned}$$

亦得矛盾. 故得 $f = \bar{f}$. 反之，若 $f = \bar{f}$，则 f 如同 $\bar{f}$ 一样凸且 lsc.

(ii) 显然 $f \leqslant \varphi \Rightarrow f^* \geqslant \varphi^*$，$\bar{f} \leqslant \bar{\varphi}$. 于是，当 φ 凸、lsc 且 $\bar{f} \leqslant \varphi \leqslant f$ 时有 $\bar{f} \leqslant \varphi = \bar{\varphi} \leqslant \bar{f}$，从而 $\bar{f} = \varphi$.

(iii) 显然 $f^* \leqslant (\bar{f})^*$. 其次，$\forall u \in X^*$，有

$$\begin{aligned}
(\bar{f})^*(u) &= \sup_{x \in X} \{u(x) - \sup_{v \in X^*} [v(x) - f^*(v)]\} \\
&= \sup_{x \in X} \inf_{v \in X^*} [\langle u - v, x \rangle + f^*(v)] \\
&\leqslant \inf_{v \in X^*} (f^*(v) + \sup_{x \in X} \langle u - v, x \rangle) = f^*(u),
\end{aligned}$$

故得 $f^* = (\bar{f})^*$. 由 $\bar{f} \leqslant f$ 得出 epi $f \subset$ epi $\bar{f}$. 因 $\bar{f}$ 必凸且为 lsc，故 epi $\bar{f}$ 闭凸，因而 $\overline{\text{co}}$ epi $f \subset$ epi $\bar{f}$. 若有

$$(x_0, t_0) \in (\text{epi } \bar{f}) \setminus \overline{\text{co}} \text{ epi } f,$$

$f^* \not\equiv \infty$，则 $\bar{f}(x_0) \leqslant t_0 < f(x_0)$，如同证(i)一样可用 1.1.2 导出矛盾. 因此 $\overline{\text{co}}$ epi $f =$ epi $\bar{f}$. □

注 5.3.3 表达了 $\bar{f}$ 的直观意义：$\bar{f}$ 是"位于 f 下方"的最大 lsc 凸函数.

5.3.4 定理 设 X 自反，$f: X \to \mathbf{R}$ 严格凸且 G-可微，$\lim\limits_{\|x\| \to \infty} \|\mathrm{D}f(x)\| = \infty$. 则 f^* 亦严格凸且 G-可微，$\mathrm{D}f^* = (\mathrm{D}f)^{-1} \stackrel{\text{def}}{=\!=} T$. f^* 有以下积分表示：

$$f^*(u) = -f(T(0)) + \int_0^1 \langle u, T(tu) \rangle \mathrm{d}t, \; u \in X^*. \qquad (8)$$

证 由定理条件推出映射 $\mathrm{D}f: X \to X^*$ 严格单调且次连续(1.5.3)，因此 $\mathrm{D}f$ 必为双射(1.3.6)，从而 $T: X^* \to X$ 严格单调且次连续(1.3.5). $\forall u, v \in X, t \in \mathbf{R}$，由式(3)及 1.5.3 有

$$\begin{aligned}
f^*(u+tv) - f^*(u) &= [\langle u+tv, T(u+tv) \rangle \\
&\quad - f(T(u+tv))] - [\langle u, Tu \rangle - f(Tu)] \\
&= \langle u, T(u+tv) - Tu \rangle - [f(T(u+tv)) - f(Tu)] \\
&\quad + t\langle v, T(u+tv) \rangle \\
&\leqslant t\langle v, T(u+tv) \rangle.
\end{aligned}$$

上式两端除以 t，并分别令 $t \downarrow 0$ 与 $t \uparrow 0$，得

$$\mathrm{D}_+ f^*(u,v) \leqslant \langle v, Tu \rangle \leqslant \mathrm{D}_- f^*(u,v),$$

因此 $\mathrm{D}f^*(u,v) = \langle v, Tu \rangle$，故得 $\mathrm{D}f^* = T$. 由 T 严格单调推出 f^* 严格凸. 由式(3)有 $0 = f(T(0)) + f^*(0)$. 于是对任给 $u \in X^*$ 有

$$\begin{aligned}
f^*(u) &= f^*(0) + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f^*(tu) \mathrm{d}t \\
&= f^*(0) + \int_0^1 \langle u, \mathrm{D}f^*(tu) \rangle \mathrm{d}t \\
&= -f(T(0)) + \int_0^1 \langle u, T(tu) \rangle \mathrm{d}t.
\end{aligned}$$

□

参考文献:[25,93,177,189,238].

§5.4 Rockafellar 对偶

Rockafellar 将共轭泛函与扰动函数结合起来,建立了一个对偶理论,它蕴涵丰富的结果.

给定 $\varphi: X \times Y \to \overline{\mathbf{R}}$,定义所谓 **S 泛函**如下:

$$S(y) = \inf_{x \in X} \varphi(x,y), \quad S_*(u) = \inf_{\lambda \in Y^*} \varphi^*(u,\lambda). \tag{1}$$

其次,由 φ 诱导出一个 Lagrange 函数:

$$L(x,\lambda) = \inf_{y \in Y} [\varphi(x,y) - \lambda(y)] = -\varphi_y^*(x,\lambda), \tag{2}$$

其中 $\varphi_y^*(x,\cdot)$ 记 $\varphi(x,\cdot)$ 的共轭泛函. 式(1),(2) 关联着一对问题:

$$(\mathrm{P}): \quad \min \varphi(x,0) = \alpha, \quad x \in X;$$

$$(\mathrm{P}^*): \quad \max [-\varphi^*(0,\lambda)] = \beta, \quad \lambda \in Y^*$$

称(P*)为(P)的 **Rockafellar 对偶**;称 $\varphi(x,y)$ 为 $\varphi(x,0)$ 的**扰动函数**.

在关于 φ 的适当假定下,可验知以下结论:

1° $S_*(u)$ 是凸函数;

$$S^*(\lambda) = \varphi^*(0,\lambda), \quad \overline{S}(0) = -S_*(0) = \beta \leqslant \alpha = S(0). \tag{3}$$

若 φ 凸,则 $S(y)$ 亦为凸函数.

2° $L(x,\cdot)$ 凹且为 usc;若 φ 凸[lsc],则 $L(\cdot,\lambda)$ 凸[lsc],从而 $L(x,\lambda)$ 为鞍[半连续]函数.

3° 注意到

$$\begin{aligned} -\varphi^*(0,\lambda) &= -\sup\{\lambda(y) - \varphi(x,y) : (x,y) \in X \times Y\} \\ &= \inf_{x \in X} \inf_{y \in Y} [\varphi(x,y) - \lambda(y)] = \inf_{x \in X} L(x,\lambda), \end{aligned}$$

故问题(P*)可表成

$$\max \inf L(X,\lambda) = \beta, \ \lambda \in Y^*.$$

其次，若 $\varphi(x,\cdot)$ 凸且为 lsc，则由式(2)与 5.3.3 有

$$\varphi(x,0)=\varphi_y^{**}(x,0)=\sup_{\lambda\in Y^*}[-\varphi_y^*(x,\lambda)]$$
$$=\sup_{\lambda\in Y^*}L(x,\lambda),$$

从而问题(P)可表成

$$\min\sup L(x,Y^*)=\alpha,\ x\in X.$$

可见当 $\varphi(x,\cdot)$ 凸且为 lsc 时，问题(P) 与(P^*) 在 §5.1 的意义下互为对偶.

4° 若 X,Y 为自反 B-空间，φ 凸且为 lsc，则互换 φ 与 φ^* 知问题(P) 与(P^*) 互为 Rockafellar 对偶.

5.4.1 定义 若 $\partial S(0)\neq\varnothing$，则说问题(P)稳定；若 $\partial S_*(0)\neq\varnothing$，则说问题($P^*$) 稳定.

利用稳定性概念，在问题(P) 与(P^*) 之间建立了异常和谐的联系. 以下是主要结果.

5.4.2 定理(Rockafellar,1967) (i) $\bar{x}$ 与 $\bar{\lambda}$ 分别为问题(P) 与(P^*) 的解 $\Leftrightarrow\alpha=\beta\neq\pm\infty\Leftrightarrow$

$$\varphi(\bar{x},0)+\varphi^*(0,\bar{\lambda})=0 \tag{4}$$

$\Leftrightarrow(0,\bar{\lambda})\in\partial\varphi(\bar{x},0)$($\Leftrightarrow(\bar{x},\bar{\lambda})$ 是 $L(x,\lambda)$ 在 $X\times Y^*$ 上的鞍点，若 $\varphi(x,y)$ 对 y 凸且 lsc).

(ii) 若 α 有限，则问题(P) 稳定 $\Leftrightarrow$ 问题(P^*) 有解且 $\alpha=\beta\Rightarrow\partial S(0)$ 是问题(P^*) 的解集；若 X,Y 自反，φ 凸且 lsc，β 有限，则问题(P^*)稳定 $\Leftrightarrow$ 问题(P)有解且 $\alpha=\beta\Rightarrow\partial S_*(0)$ 是(P) 的解集.

(iii) 若 $\alpha>-\infty$，φ 为凸函数，存在 $\hat{x}\in X$ 与 Y 的 0- 邻域 V，使 $\sup\varphi(\hat{x},V)<\infty$，则问题(P) 稳定；若 $\beta<\infty$，存在 $\hat{\lambda}\in Y^*$ 与 X^* 的 0- 邻域 N，使 $\sup\varphi^*(N,\hat{\lambda})<\infty$，则问题($P^*$) 稳定.

证 (i) 由 Young 不等式(§5.3(2))有

$$0\leqslant\varphi(x,0)+\varphi^*(0,\lambda),(x,\lambda)\in X\times Y^*,$$

且仅当 $(0,\lambda)\in\partial\varphi(x,0)$ 时上式取等号. 这一事实结合 5.1.1 得出结论(i).

(ii) 只需证前一半. 由(3) 及 §5.3 中的结论,知 $\lambda \in \partial S(0) \Leftrightarrow 0 = S(0) + S^*(\lambda) = \alpha + \varphi^*(0,\lambda) \Leftrightarrow \alpha = -\varphi^*(0,\lambda) \Leftrightarrow \lambda$ 是问题(P^*)的解且 $\alpha = \beta$. 由此得出:$\partial S(0) \neq \varnothing \Leftrightarrow$ 问题(P^*) 有解且 $\alpha = \beta$.

(iii) 同样只需证前一半. 由 φ 凸推出 $S(y)$ 凸;由 $\sup \varphi(\hat{x},V) < \infty$ 推出 $S(y)$ 在 V 内上有界;由 $S(0) = \alpha > -\infty$ 推出 $S(y)$ 在 $y = 0$ 邻近有限(否则有 $y \approx 0: S(y) = -\infty$,从而 $2S(0) \leqslant S(y) + S(-y) = -\infty$,这矛盾于 $\alpha > -\infty$). 于是 $S(y)$ 在 $y = 0$ 连续(1.5.2),从而 $\partial S(0) \neq \varnothing$(2.1.3),即(P) 稳定. □

注 5.4.2(iii) 中的条件 $\sup \varphi(\hat{x},V) < \infty$ 可代以等价的条件"$\varphi(\hat{x},y)$ 在 $y = 0$ 连续"(注意已假定 φ 凸). 条件"$\sup \varphi^*(N,\hat{\lambda}) < \infty$"仿此.

5.4.3 推论 (i) 若 X,Y 自反,φ 凸且为 lsc,则问题(P) 稳定且有解 $\Leftrightarrow$(P^*) 稳定且有解 $\Leftrightarrow$(P) 与(P^*) 皆稳定且 α 有限 $\Leftrightarrow$ 问题(P) 与问题(P^*) 皆有解且 $\alpha = \beta \neq \infty \Leftrightarrow L(x,\lambda)$ 有鞍点.

(ii)(必要条件) 若 $\bar{x}$ 是问题(P) 的解,φ 凸且 $\exists \hat{x} \in X: \varphi(\hat{x},y)$ 在 $y = 0$ 连续,则有 $\bar{\lambda} \in Y^*$ 使条件式(4) 满足.

(iii)(充分条件) 若有$(\bar{x},\bar{\lambda}) \in X \times Y^*$ 使条件式(4) 满足,则 $\bar{x}$ 是问题(P) 的解.

(iv)(充分条件) 若 X,Y 自反,φ 凸且为 lsc,β 有限,$\exists \hat{\lambda} \in Y^*$: $\varphi^*(u,\hat{\lambda})$ 在 $u = 0$ 连续,则问题(P) 有解且 $\alpha = \beta$.

以上结论可与 §4.3 的结果对照. 条件式(4) 与 K-T 条件相当,在某些场合它确能转化为 K-T 条件. 条件"$\exists \hat{x} \in X: \varphi(\hat{x},y)$ 在 $y = 0$ 连续"相当于 Slater 条件(3.4.2 条件(iv)),在一定情况下它确与 Slater 条件一致. 此条件有点强,不少新近研究致力于减弱此条件(参看 5.6.6). 值得注意的是,5.4.2 与 5.4.3 中的条件完全不涉及导数.

若 $\dim Y < \infty$,凸函数 $S(y)$ 在 $y = 0$ 邻近有限,则 $S(y)$ 必在 $y = 0$ 连续(1.5.2). 因此有

5.4.4 推论 设 $\dim Y < \infty$,φ 凸,$\alpha > -\infty$,在 $y = 0$ 邻近

$S(y) < \infty$，则问题(P)稳定，且问题(P*)有解.

在 Rockafellar 理论的具体应用中，出发点通常不是某个给定的 φ，而是某个给定的"原问题"(P). 应用 Rockafellar 理论的关键在于引进适当的**扰动变量** y 并构成**扰动函数** $\varphi(x,y)$，使得原问题等价于问题"$\min \varphi(x,0) = \alpha, x \in X$"，从而得到其 Rockafellar 对偶(P*). φ 的构成与原问题的目标函数及约束的特性有关，是一个颇带技术性的问题. 一些典型的方法将在下节考虑.

参考文献：[16,25,93,158,238].

§5.5 Fenchel 对偶：一般情况

给定 $\omega: X \to \overline{\mathbf{R}}, \gamma: Y \to \overline{\mathbf{R}}, \eta: Z \to \overline{\mathbf{R}}, g: X \to Y, h: X \to Z$，考虑如下无约束极小问题

$$\min\left[\omega(x) + \gamma(g(x)) + \eta(h(x))\right] = \alpha, x \in X. \tag{1}$$

为应用 Rockafellar 理论，构成如下扰动函数：

$$\varphi(x,y,z) = \omega(x) + \gamma(g(x) + y) + \eta(h(x) + z). \tag{2}$$

于是问题(1)可改写成

$$\min \varphi(x,0,0) = \alpha, \ x \in X.$$

在关于 ω,γ,η,g,h 的适当假定下，可验知以下结论：

1° 若 γ 与 η 凸[lsc]，则 φ 对 (y,z) 为凸[lsc]；若 ω,γ,η 凸且 γ 依 Y_+ 导出的序单调增，g 为 Y_+ - 凸，h 是仿射函数，则 φ 凸.

2° 任给 $u \in X^*, \lambda \in Y^*, \mu \in Z^*$，有

$$\begin{aligned}\varphi^*(u,\lambda,\mu) &= \sup\{u(x) + \lambda(y) + \mu(z) - \varphi(x,y,z) : \\ &\qquad (x,y,z) \in X \times Y \times Z\} \\ &= \sup_{x,y,z}[u(x) + \lambda(y - g(x)) + \mu(z - h(x)) \\ &\qquad - \omega(x) - \gamma(y) - \eta(z)] \\ &= \sup_{x,y,z}[u(x) - \Gamma(x,\lambda,\mu) + \lambda(y) - \gamma(y) + \mu(z) - \eta(z)]\end{aligned}$$

$$= \Gamma_x^*(u,\lambda,\mu) + \gamma^*(\lambda) + \eta^*(\mu), \tag{3}$$

其中 $\Gamma(\cdot,\lambda,\mu) = \omega + \lambda g + \mu h$；$\Gamma_x^*$ 记 $\Gamma(x,\lambda,\mu)$ 对 x 的共轭函数. 于是，问题(1) 的 Rockafellar 对偶为

$$\begin{cases} \max[-\Gamma_x^*(0,\lambda,\mu) - \gamma^*(\lambda) - \eta^*(\mu)] = \beta, \\ (\lambda,\mu) \in Y^* \times Z^*. \end{cases} \tag{1*}$$

称(1*)为问题(1)的 **Fenchel 对偶**(在推广的意义上).

3° 结合函数(2)与函数(3)有

$$\begin{aligned} &\varphi(x,0,0) + \varphi^*(0,\lambda,\mu) \\ &\quad = [\Gamma(x,\lambda,\mu) + \Gamma_x^*(0,\lambda,\mu) - \langle 0,x\rangle] \\ &\qquad + [\gamma(g(x)) + \gamma^*(\lambda) - \langle\lambda,g(x)\rangle] \\ &\qquad + [\eta(h(x)) + \eta^*(\mu) - \langle\mu,h(x)\rangle]. \end{aligned}$$

由 Young 不等式，上式右端三项皆非负. 因此，条件 $\varphi(\bar{x},0,0) + \varphi^*(0,\bar{\lambda},\bar{\mu}) = 0$ (参照 § 5.4 之式(4))满足的充要条件是

$$\begin{cases} \Gamma(\bar{x},\bar{\lambda},\bar{\mu}) + \Gamma_x^*(0,\bar{\lambda},\bar{\mu}) = 0, \\ \gamma(g(\bar{x})) + \gamma^*(\bar{\lambda}) = \bar{\lambda}g(\bar{x}), \\ \eta(h(\bar{x})) + \eta^*(\bar{\mu}) = \bar{\mu}h(\bar{x}); \end{cases}$$

而这又等价于(参考 § 5.3)

$$0 \in \partial_x\Gamma(\bar{x},\bar{\lambda},\bar{\mu}), \bar{\lambda} \in \partial\gamma(g(\bar{x})), \bar{\mu} \in \partial\eta(h(\bar{x})). \tag{4}$$

其中 $0 \in \partial_x\Gamma(\bar{x},\bar{\lambda},\bar{\mu}) \Leftrightarrow \Gamma(x,\bar{\lambda},\bar{\mu}) = \min \Gamma(X,\bar{\lambda},\bar{\mu})$.

4° 若令 $\Lambda(x,\lambda,\mu) = -\varphi^*_{(y,z)}(x,\lambda,\mu)$(参照 § 5.4 之式(2))，则类似于 2° 中的计算可得

$$\Lambda(x,\lambda,\mu) = \begin{cases} \Gamma(x,\lambda,\mu) - \gamma^*(\lambda) - \eta^*(\mu), x \in D_\omega, \\ \infty, \quad \text{否则}. \end{cases} \tag{5}$$

5° 若 $\hat{x} \in D_\omega$，γ 与 η 分别在 $y = g(\hat{x})$ 与 $z = h(\hat{x})$ 连续，则 $\varphi(\hat{x},y,z)$ 在 $(y,z) = (0,0)$ 连续. 若 $\gamma^*(\hat{\lambda})$ 与 $\eta^*(\hat{\mu})$ 有限，$\Gamma_x^*(u,\hat{\lambda},\hat{\mu})$ 在 $u = 0$ 连续，则 $\varphi^*(u,\hat{\lambda},\hat{\mu})$ 在 $u = 0$ 连续. 若 $\omega(\hat{x})$，$\gamma(g(\hat{x}))$ 与 $\eta(h(\hat{x}))$ 有限，则 $\alpha < \infty$；若 $\Gamma_x^*(0,\hat{\lambda},\hat{\mu})$，$\gamma^*(\hat{\lambda})$ 与 $\eta^*(\hat{\mu})$ 有限，则 $\beta > -\infty$.

结合以上结论与 5.4.2 得出：

5.5.1 定理　对问题(1) 与(1^*) 以下结论成立：

(i) $\bar{x}$ 与 $(\bar{\lambda},\bar{\mu})$ 分别为问题(1) 与(1^*) 的解且 $\alpha=\beta\neq\pm\infty\Leftrightarrow$ 条件(4) 满足($\Leftrightarrow(\bar{x},\bar{\lambda},\bar{\mu})$ 为 $\Lambda(x,\lambda,\mu)$ 在 $X\times(Y^*\times Z^*)$ 上的鞍点，若 γ,η 凸且 lsc).

(ii) 若 α 有限，则问题(1) 稳定 $\Leftrightarrow$ 问题(1^*) 有解且 $\alpha=\beta$；若 X,Y,Z 自反，ω,γ,η 凸且 γ 单调增，g 为 Y_+-凸，h 为仿射函数，φ 为 lsc，β 有限，则问题(1^*) 稳定 $\Leftrightarrow$ 问题(1) 有解且 $\alpha=\beta$.

(iii) 若 $\alpha>-\infty$，ω,γ,η 凸且 γ 单调增，g 为 Y_+-凸，h 为仿射函数，$\exists\hat{x}\in D_\omega$：$\gamma$ 与 η 分别在 $g(\hat{x})$ 与 $h(\hat{x})$ 邻近上有界，则问题(1) 稳定；若 $\beta<\infty$，$\exists\hat{\lambda}\in Y^*,\hat{\mu}\in Z^*$：$\gamma^*(\hat{\lambda})$ 与 $\eta^*(\hat{\mu})$ 有限且函数 $\Gamma_x^*(u,\hat{\lambda},\hat{\mu})$ 在 $u=0$ 连续，则问题(1^*) 稳定.

5.5.1 有很大的一般性，其具体应用取决于函数 ω,γ,η,g,h 的特殊选择. 下面指明，适当地选取 ω,γ,η，可从 Rockafellar 对偶导出 Lagrange 对偶.

取 $\omega=f+\delta_D,\gamma=\delta_{Y_-},\eta=\delta_{\{0\}}(0\in Z)$，则问题(1) 与扰动函数 φ 依次成为

$$\min f(x)=\alpha,g(x)\leqslant 0,h(x)=0,x\in D \tag{6}$$

与
$$\varphi(x,y,z)=\begin{cases}f(x),g(x)\leqslant -y,h(x)=-z,x\in D,\\ \infty, & \text{否则}.\end{cases} \tag{7}$$

令 $L(\cdot,\lambda,\mu)=f+\lambda g+\mu h$，则 $\Gamma=L+\delta_D$. 因 $\gamma^*=\delta_{Y_+^*},\eta^*=0$(5.3.2)，$g(\bar{x})\leqslant 0\Rightarrow\partial\gamma(g(\bar{x}))=(g(x)+Y_+)^*=Q=Y_+^*\cap\{g(\bar{x})\}^\perp$，$\partial\eta(0)=Z^*$(2.1.2)，故在上述具体情况下式(3)，(4)，(5)依次成为

$$\varphi^*(u,\lambda,\mu)=\begin{cases}\sup\limits_{x\in D}\ [u(x)-L(x,\lambda,\mu)],\lambda\in Y_+^*,\\ \infty, & \text{否则};\end{cases} \tag{8}$$

$$f(\bar{x})=\min L(D,\bar{\lambda},\bar{\mu}),(\bar{x},\bar{\lambda},\bar{\mu})\in M\times Q\times Z^*; \tag{9}$$

$$
\Lambda(x,\lambda,\mu) = \begin{cases} L(x,\lambda,\mu), x \in D, \lambda \in Y_+^*, \\ -\infty, \quad x \in D, \lambda \overline{\in} Y_+^*, \\ \infty \quad x \overline{\in} D. \end{cases} \tag{10}
$$

由函数(8)知问题(6)的 Rockafellar 对偶为

$$
\max \inf L(D,\lambda,\mu) = \beta, (\lambda,\mu) \in Y_+^* \times Z^*, \tag{6*}
$$

这正是问题(6)的 Lagrange 对偶(参看 § 5.2(1*)). 此外,直接看出 $\gamma = \delta_{Y_-}$ 与 $\eta = \delta_{\{0\}}$ 凸且为 lsc, γ 单调增. 于是结合以上事实与 5.5.1 得出:

5.5.2 推论　对问题(6)与(6*)以下结论成立:

(i) $\bar{x}$ 与 $(\bar{\lambda},\bar{\mu})$ 分别为问题(6)与(6*)的解且 $\alpha = \beta \neq \pm\infty \Leftrightarrow$ 条件式(9)(它与 § 5.2(2) 一致)满足 $\Leftrightarrow (\bar{x},\bar{\lambda},\bar{\mu})$ 是 $L(x,\lambda,\mu)$ 在 $D \times (Y_+^* \times Z^*)$ 上的鞍点.

(ii) 若 α 有限,则问题(6)稳定 $\Leftrightarrow$ 问题(6*)有解且 $\alpha = \beta$; 若 X, Y, Z 自反, f 凸且为 lsc, g 为 Y_+ - 凸且为 * lsc, D 闭凸, h 为仿射函数, β 有限,则问题(6*)稳定 $\Leftrightarrow$ 问题(6)有解且 $\alpha = \beta$.

(iii) 若 $\beta < \infty$, 存在 $(\hat{\lambda},\hat{\mu}) \in Y_+^* \times Z^*$, 使函数 $u \to \sup_{x \in D}[u(x) - L(x,\hat{\lambda},\hat{\mu})]$ 在 $u = 0$ 连续,则问题(6*)稳定.

注意 5.5.2(i) 与 5.2.1(i) 完全一致.

若取 $Z = \{0\}$, 则以上讨论有所简化. 此时问题(1)成为

$$
\min [\omega(x) + \gamma(g(x))] = \alpha, \; x \in X. \tag{11}
$$

相应地,式(2)~式(5)分别简化为

$$
\varphi(x,y) = \omega(x) + \gamma(g(x) + y); \tag{12}
$$

$$
\varphi^*(u,\lambda) = \Gamma_x^*(u,\lambda) + \gamma^*(\lambda); \tag{13}
$$

$$
0 \in \partial_x \Gamma(\bar{x},\bar{\lambda}), \lambda \in \partial\gamma(g(\bar{x})); \tag{14}
$$

$$
\Lambda(x,\lambda) = \begin{cases} \Gamma(x,\lambda) - \gamma^*(\lambda), x \in D_\omega, \\ \infty, \quad \text{否则}, \end{cases} \tag{15}
$$

其中 $\Gamma(\cdot,\lambda) = \omega + \lambda g$,

$$
0 \in \partial_x \Gamma(\bar{x},\bar{\lambda}) \Leftrightarrow \Gamma(\bar{x},\bar{\lambda}) = \min \Gamma(X,\bar{\lambda})
$$

$$\Leftrightarrow \Gamma(\bar{x},\bar{\lambda}) + \Gamma_x^*(0,\bar{\lambda}) = 0.$$

问题(11)的 Fenchel 对偶是(参照(1*)):

$$\max[-\Gamma_x^*(0,\lambda) - \gamma^*(\lambda)] = \beta,\ \lambda \in Y^*. \qquad (11^*)$$

利用以上结论,从 5.5.1 得出:

5.5.3 推论 对问题(11)与(11*)以下结论成立:

(i) $\bar{x}$ 与 $\bar{\lambda}$ 分别为问题(11) 与(11*) 的解且 $\alpha = \beta \neq \pm\infty$ $\Leftrightarrow$ 条件(14) 满足($\Leftrightarrow(\bar{x},\bar{\lambda})$ 是 $\Lambda(x,\lambda)$ 的鞍点,若 γ 凸且 lsc).

(ii) 若 α 有限,则问题(11) 稳定 $\Leftrightarrow$ 问题(11*) 有解且 $\alpha = \beta$; 若 X,Y 自反,ω,γ,g 凸,γ 单调增,φ 为 lsc,β 有限,则问题(11*) 稳定 $\Leftrightarrow$ 问题(11) 有解且 $\alpha = \beta$.

(iii) 若 $\alpha > -\infty$,ω,γ 凸且 γ 单调增,g 为 Y_+-凸,$\exists \hat{x} \in D_\omega$: γ 在 $g(\hat{x})$ 邻近上有界,则问题(11) 稳定;若 $\beta < \infty$,$\exists \hat{\lambda} \in Y^*$: $\gamma^*(\hat{\lambda})$ 有限,$\Gamma_x^*(u,\hat{\lambda})$ 在 $u = 0$ 连续,则问题(11*) 稳定.

将 5.5.3 用于问题

$$\min f(x) = \alpha,\ g(x) \leqslant 0,\ x \in D \qquad (16)$$

与其 Lagrange 对偶(参看 §5.2(5*)):

$$\max \inf L(D,\lambda) = \beta,\ \lambda \in Y_+^*, \qquad (16^*)$$

其中 $L(\cdot,\lambda) = f + \lambda g$,得到对应于 5.5.2 的以下结论.

5.5.4 推论 (i) $\bar{x}$ 与 $\bar{\lambda}$ 分别为问题(16) 与(16*) 的解且 $\alpha = \beta \neq \pm\infty \Leftrightarrow (\bar{x},\bar{\lambda}) \in M \times Q$ 且 $f(\bar{x}) = \min L(D,\bar{\lambda})$($M = D \cap g^{-1}(Y_-)$)$\Leftrightarrow(\bar{x},\bar{\lambda})$ 是 $L(x,\lambda)$ 在 $D \times Y_+^*$ 上的鞍点.

(ii) 若 α 有限,则问题(16) 稳定 $\Leftrightarrow$ 问题(16*) 有解且 $\alpha = \beta$; 若 X,Y 自反,f 凸且为 lsc,g 为 Y_+-凸且为 $*$ lsc,D 闭凸,β 有限,则问题(16*) 稳定 $\Leftrightarrow$ 问题(16) 有解且 $\alpha = \beta$.

(iii) 若 $\alpha > -\infty$,$f|D$ 凸,g 为 Y_+-凸,$\exists \hat{x} \in D$: $g(\hat{x}) \ll 0$,则问题(16) 稳定;若 $\beta < \infty$,存在 $\hat{\lambda} \in Y_+^*$,使函数 $u \to \sup_{x \in D}[u(x) - L(x,\hat{\lambda})]$ 在 $u = 0$ 连续,则问题(16*) 稳定.

关于 Fenchel 对偶与 Lagrange 对偶的关系的更深入的讨论,

参看文献[51],例如[51,Th. 2].

§5.6 Fenchel 对偶:特殊情况

本节考虑上节问题(11)的特款

$$\min\,[f(x) + \gamma(Ax - a)] = \alpha,\ x \in D, \tag{1}$$

其中 $\gamma: Y \to \overline{\mathbf{R}}$, $A \in L(X,Y)$, $a \in Y$. 若令 $\omega = f + \delta_D$, $g(x) = Ax - a$, 则问题(1)可写成形如 §5.5 中式(11)的标准形式.

为应用 5.5.3,首先指明以下结论:

1° 由 §5.5 中式(12),扰动函数

$$\varphi(x,y) = f(x) + \delta_D(x) + \gamma(Ax - a + y). \tag{2}$$

若 γ 凸[lsc],则 $\varphi(x,\cdot)$ 凸[lsc];若 D 闭凸,f 与 γ 凸[lsc],则 φ 凸[lsc].

2° 令 $L(x,\lambda) = f(x) + \langle\lambda, Ax - a\rangle$, $\Gamma(x,\lambda) = L(x,\lambda) + \delta_D(x)$,则

$$\begin{aligned}\Gamma_x^*(u,\lambda) &= \sup_{x\in D}\,[u(x) - f(x) \quad \langle\lambda, Ax - a\rangle]\\ &= \sup_{x\in D}\,[\langle u - A^*\lambda, x\rangle - f(x)] + \lambda(a)\\ &\quad (= f^*(u - A^*\lambda) + \lambda(a),\text{若 } D = X).\end{aligned}$$

于是由 §5.5 之式(13)有

$$\begin{aligned}\varphi^*(u,\lambda) &= \sup_{x\in D}[\langle u - A^*\lambda, x\rangle - f(x)] + \gamma^*(\lambda) + \lambda(a) \quad (3)\\ &\quad (= f^*(u - A^*\lambda) + \gamma^*(\lambda) + \lambda(a),\text{若 } D = X).\end{aligned}$$

因此,问题(1)的 Fenchel 对偶为

$$\max\,\{\inf_{x\in D}[\langle A^*\lambda, x\rangle + f(x)] - \gamma^*(\lambda) - \lambda(a)\} = \beta,\ \lambda \in Y^*. \tag{1*}$$

当 $D = X$ 时问题(1*)简化为

$$\max[-f^*(-A^*\lambda) - \gamma^*(\lambda) - \lambda(a)] = \beta,\ \lambda \in Y^*.$$

3° 由§5.5之式(14)，$\varphi(\bar{x},0)+\varphi^*(0,\bar{\lambda})=0$ 等价于

$$L(\bar{x},\bar{\lambda})=\min L(D,\bar{\lambda}),\bar{\lambda}\in\partial\gamma(A\bar{x}-a),\bar{x}\in D. \qquad (4)$$

若 $D=X$，则 $L(\bar{x},\bar{\lambda})=\min L(D,\bar{\lambda})\Leftrightarrow -A^*\bar{\lambda}\in\partial f(\bar{x})$.

4° 依§5.5之式(15)，有

$$\Lambda(x,\lambda)=\begin{cases}L(x,\lambda)-\gamma^*(\lambda), & x\in D,\\ \infty, & \text{否则}.\end{cases} \qquad (5)$$

结合上述与5.5.3得(参考[51,Th.1])：

5.6.1 定理 对问题(1)与(1*)以下结论成立：

(i) $\bar{x}$与$\bar{\lambda}$分别为问题(1)与(1*)的解且$\alpha=\beta\neq\pm\infty\Leftrightarrow$式(4)成立($\Leftrightarrow(\bar{x},\bar{\lambda})$是$\Lambda(x,\lambda)$在$D\times Y^*$上的鞍点，若$\gamma$凸且为lsc).

(ii) 若α有限，则(1)稳定$\Leftrightarrow$问题(1*)有解且$\alpha=\beta$；若X,Y自反，f与γ凸且为lsc，D闭凸，β有限，则问题(1*)稳定$\Leftrightarrow$问题(1)有解且$\alpha=\beta$.

(iii) 若$\alpha>-\infty$，f与γ凸，$\exists\hat{x}\in D$：$\gamma(y)$在$y=A\hat{x}-a$连续，则问题(1)稳定；若$\beta<\infty$，$\exists\hat{\lambda}$：$\gamma^*(\hat{\lambda})$有限，函数$u\to\sup\limits_{x\in D}[\langle u-A^*\hat{\lambda},x\rangle-f(x)]$在$u=0$连续，则问题(1*)稳定.

进一步的结果由取更特殊的f或$g(x)=Ax-a$得出. 首先，设$f\in X^*$，考虑问题

$$\min[f(x)+\gamma(Ax-a)]=\alpha,\ x\in X. \qquad (6)$$

因$f^*=\delta_{\{f\}}(5.3.2)$，故问题(6)的Fenchel对偶为

$$\max[-\gamma^*(\lambda)-\lambda(a)]=\beta,A^*\lambda+f=0,\lambda\in Y^*. \quad (6^*)$$

若γ是G-可微的，则条件(4)可写成：

$$A^*\bar{\lambda}+f=0,\bar{\lambda}=D\gamma(A\bar{x}-a). \qquad (7)$$

由式(7)推出$A^*D\gamma(A\bar{x}-a)+f=0$. 结合上述与5.6.1及5.3.4得出：

5.6.2 定理 设$f\in X^*$，$\gamma: Y\to\mathbf{R}$严格凸且G-可微.

(i) $\bar{x}$与$\bar{\lambda}$分别为问题(6)与(6*)的解且$\alpha=\beta\Leftrightarrow$条件(7)成立$\Leftrightarrow(\bar{x},\bar{\lambda})$是$L(x,\lambda)-\gamma^*(\lambda)$在$X\times Y^*$上的鞍点.

(ii) 若X自反，$\lim\limits_{\|y\|\to\infty}\|D\gamma(y)\|=\infty$，存在$c>0$使$\|Ax\|\geqslant$

$c|x|$，则问题(6)与(6*)分别有唯一解 $\bar{x}$ 与 $\bar{\lambda}$，且 $(\bar{x},\bar{\lambda})$ 是 $L(x,\lambda) - \gamma^*(\lambda)$ 的鞍点.

证 只需证(ii). 由 γ 严格凸推出 $D\gamma$ 严格单调(1.5.3)，从而 $A^* D\gamma(Ax-a)+f$ 亦严格单调. 易见 $|x| \to \infty \Rightarrow |Ax-a| \to \infty \Rightarrow |A^* D\gamma(Ax-a)+f| \to \infty$，故有唯一 $\bar{x} \in X$，使

$$A^* D\gamma(A\bar{x}-a)+f=0$$

(参考 1.3.6). 令 $\bar{\lambda} = D\gamma(A\bar{x}-a)$，则 $(\bar{x},\bar{\lambda})$ 满足条件式(7)，因此 $\bar{x}$ 与 λ 分别为问题(6)与(6*)的解. 因 γ 与 γ^* 皆严格凸(5.3.4)，故唯一性结论成立. □

应用 5.6.1 的另一途径是保持 f 的一般性，但特取 $D=X=Y$，$Ax-a \equiv x$，即考虑问题

$$\min\,[f(x)+\gamma(x)]=\alpha,\ x \in X. \tag{8}$$

称问题(8)为**和极小问题**，关于它有一系列研究. 由问题(1*)得出问题(8)的 Fenchel 对偶为

$$\max\,[-f^*(-u)-\gamma^*(u)]=\beta,\ u \in X^*. \tag{8*}$$

直接应用 5.6.1 得：

5.6.3 推论 (i) $\bar{x}$ 与 $\bar{u}$ 分别为问题(8)与(8*)的解且 $\alpha=\beta \neq \pm\infty \Leftrightarrow \bar{u} \in (-\partial f(\bar{x})) \cap \partial\gamma(\bar{x})$ ($\Leftrightarrow (\bar{x},\bar{u})$ 是 $\Lambda(x,u)=f(x)+u(x)-\gamma^*(u)$ 的鞍点，若 γ 凸且为 lsc).

(ii) 若 α 有限，则问题(8)稳定 $\Leftrightarrow$ 问题(8*)有解且 $\alpha=\beta$；若 X 自反，f 与 γ 凸且为 lsc，β 有限，则问题(8*)稳定 $\Leftrightarrow$ 问题(8)有解且 $\alpha=\beta$.

(iii) 若 $\alpha>-\infty$，f 与 γ 凸，$\exists \hat{x} \in D_f$：γ 在 $\hat{x}$ 连续，则问题(8)稳定；若 $\beta<\infty$，$\exists \hat{u}$：$\gamma^*(\hat{u})$ 有限，$f^*(u)$ 在 $u=-\hat{u}$ 连续，则问题(8*)稳定.

更深入的结论依赖于某些新的概念.

5.6.4 定义 设 $D \subset X$. 令

$$\mathrm{ri}\,D=\{x \in D:\mathrm{cone}\,(D-x)\ \text{为子空间}\};$$

$$\mathrm{qri}\,D=\{x \in D:\overline{\mathrm{cone}}(D-x)\ \text{为子空间}\};$$

$$\mathrm{sqri}\,D=\{x \in D:\mathrm{cone}\,(D-x)\ \text{为闭子空间}\}.$$

ri D, qri D 与 sqri D 分别称为 D 的相对内部、拟相对内部与强拟相对内部.

若 dim $X<\infty, D\subset X$,则易验证 ri $D=$ qri $D=$ sqri D.

5.6.5 定理 假定:(i) f,γ 凸且为 lsc; (ii) $0\in$ spri (D_f-D_γ). 若问题(8) 的最优值 α 有限,则问题(8*) 有解且 $\alpha=\beta$.

证 1° 设 $0\in D_f\cap D_\gamma$. 令 $X_0=\text{cone}\,(D_f-D_\gamma)$,则 X_0 是 X 的闭子空间,$D_f,D_\gamma\subset X_0$. 当 $X_0=X$ 时,定理结论由 Rockafellar 的一个结果推出,这种情况的证明从略. 若 $X_0\neq X$,则令 $f_0=f|X_0,\gamma_0=\gamma|X_0$,对 f_0,γ_0 在 X_0 上应用定理结论知有 $u_0\in X_0^*$(由 1.1.1 可设 $u_0\in X^*$),使得

$$
\begin{aligned}
\alpha &\leqslant \inf\{f(x)+\gamma(x): x\in X_0\}\\
&=\max\{-f_0^*(-u)-\gamma_0^*(u): u\in X_0^*\}\\
&=-f_0^*(-u_0)-\gamma_0^*(u_0)\\
&=-f^*(-u_0)-\gamma^*(u_0)\leqslant\beta
\end{aligned}
$$

(注意 $f_0^*(u)=f^*(u)$),由此得出定理结论.

2° 设 $0\,\overline{\in}\,D_f\cap D_\gamma$. 由条件(ii) 有 $0\in D_f-D_\gamma$,故有 $x_0\in D_f\cap D_\gamma$. 令 $f_1(x)=f(x+x_0),\gamma_1(x)=\gamma(x+x_0)$,则 $0\in D_{f_1}\cap D_{\gamma_1}$. 于是由已证的 1° 有

$$
\begin{aligned}
\alpha &= \inf\{f_1(x)+\gamma_1(x): x\in X\}\\
&=\max\{-f_1^*(-u)-\gamma_1^*(u): u\in X^*\}\\
&=\max\{-f^*(-u)-\gamma^*(u): u\in X^*\}
\end{aligned}
$$

(利用 § 5.3(5)),由此得出定理结论. □

注 若 f,γ 凸且为 lsc,α 有限,γ 在某点 $\hat{x}\in D_f$ 邻近上有界,则 $\hat{x}\in D_f\cap D_\gamma^\circ$,从而 $0\in(D_f-D_\gamma)^\circ\subset$ sqri (D_f-D_γ). 于是由 5.6.5 知问题(8*)有解且 $\alpha=\beta$. 这一结论已包含于 5.6.3. 注意条件 $0\in$ sqri (D_f-D_γ) 远宽于条件 $D_f\cap D_\gamma^\circ\neq\varnothing$!

5.6.5 可用来改进 5.4.2 中的某些结论. 设 $\varphi: X\times Y\to\overline{\mathbf{R}}$,考虑一对问题:

$$\min \varphi(x,0) = \alpha,\ x \in X; \tag{9}$$

$$\max [-\varphi^*(0,\lambda)] = \beta,\ \lambda \in Y^*. \tag{9*}$$

5.6.6 定理 假定:(i) φ凸且为lsc; (ii) $0 \in \operatorname{sqri}(\pi D_\varphi)$,其中$\pi: X \times Y \to Y$是投影.若问题(9)的最优值$\alpha$有限,则问题(9*)有解且$\alpha = \beta$.

证 令$\psi = \delta_{X\times 0}$,则$\psi$凸且为lsc.由条件(ii)知cone (πD_φ)是Y的闭子空间,从而

$$\begin{aligned}\operatorname{cone}(D_\varphi - D_\psi) &= \operatorname{cone}(D_\varphi - X \times 0)\\ &= \operatorname{cone} X \times (\pi D_\varphi) = X \times \operatorname{cone}(\pi D_\varphi)\end{aligned}$$

是$X \times Y$的闭子空间,即$0 \in \operatorname{sqri}(D_\varphi - D_\psi)$.于是以$\varphi,\psi$代$f,\gamma$从5.6.5得出:

$$\begin{aligned}\alpha &= \inf \varphi(X,0)\\ &= \inf\{\varphi(x,y) + \psi(x,y) : (x,y) \in X \times Y\}\\ &= \max\{-\varphi^*(u,\lambda) - \psi^*(-u,-\lambda) : (u,\lambda) \in X^* \times Y^*\}\\ &= \max[-\varphi^*(0,Y^*)],\end{aligned}$$

其中用到$\psi^* = s_{X\times 0}$(5.3.2之5°).由此得出定理结论. □

关于5.6.5与5.6.6可参看[109].

注 若φ凸且lsc,α有限,$\exists \hat{x} \in X$:$\varphi(\hat{x},y)$在$y=0$邻近上有界,则$0 \in (\pi D_\varphi)^\circ \subset \operatorname{sqri}(\pi D_\varphi)$.于是由5.6.6知问题(9*)有解且$\alpha = \beta$.这一结论已包含于5.4.2.

将5.6.6用到

$$\varphi(x,y) = \begin{cases} f(x), x \in D \text{ 且 } g(x) + y \leqslant 0,\\ \infty, \quad \text{否则}, \end{cases}$$

得出关于Lagrange对偶的以下结论:

5.6.7 推论 假定:(i) $f|D$凸且为lsc,$g|D$为Y_+-凸且$*$lsc,D为闭凸;(ii) cone $(g(D) + Y_+)$为Y的闭子空间.若$\alpha = \inf\{f(x) : x \in D \text{ 且 } g(x) \leqslant 0\}$有限,$L(\cdot,\lambda) = f + \lambda g$,则

$$\alpha = \max_{\lambda \in Y_+^*} \inf L(D,\lambda).$$

注 若$\exists \hat{x} \in D$:$g(\hat{x}) \ll 0$,则$0 \in (g(D) + Y_+)^\circ$,从而得出

cone $(g(D)+Y_+)=Y$,5.6.7 之条件(ii) 满足(参看 5.5.4).

参考文献:[1,51,93,109,151,158,238].

§ 5.7 Mond-Weir 对偶与 Wolfe 对偶

设 f,g,h 是可微函数. 对于问题

$$\min f(x)=\alpha,\quad g(x)\leqslant 0,\ h(x)=0, \tag{1}$$

所谓 **Mond-Weir 对偶与 Wolfe 对偶**分别为

$$\begin{cases}\max f(x)=\beta,\\ L_x(x,\lambda,\mu)=0,\lambda g(x)+\mu h(x)\geqslant 0,\\ (\lambda,\mu)\in Y_+^*\times Z^*;\end{cases} \tag{2}$$

$$\begin{cases}\max L(x,\lambda,\mu)=\beta,\\ L_x(x,\lambda,\mu)=0,(\lambda,\mu)\in Y_+^*\times Z^*.\end{cases} \tag{3}$$

涉及以上两种对偶的文献甚多,大多用到一定的凸性条件.

以下分别以 M,N,P 记问题(1),(2),(3)的可行集. 若 $\alpha\geqslant\beta$,则说问题(1) 与(2)(或问题(1) 与(3)) **弱对偶**. 若问题(1) 与(2) 弱对偶,$x\in M,(\bar{x},\bar{\lambda},\bar{\mu})\in N$,则 $\bar{x}$ 与$(\bar{x},\bar{\lambda},\bar{\mu})$ 分别为问题(1) 与(2) 的最优解,且 $\bar{x}$ 是(1) 的 K-T 点,$\bar{\lambda},\bar{\mu}$ 为 K-T 乘子. 若问题(1) 与(3) 弱对偶,$\bar{x}\in M,(\bar{x},\bar{\lambda},\bar{\mu})\in P,\bar{\lambda}g(\bar{x})=0$,则 $\bar{x}$ 与$(\bar{x},\bar{\lambda},\bar{\mu})$ 分别为问题(1) 与(3) 的最优解,且 $\bar{x}$ 是(1) 的 K-T 点,$\bar{\lambda},\bar{\mu}$ 是 K-T 乘子. 可见,给出弱对偶的条件有重要意义.

5.7.1 定理(弱对偶性) 以下每个条件推出问题(1) 与(2) 弱对偶:

(i) $\forall(x,\lambda,\mu)\in N,L(\cdot,\lambda,\mu)$ 在 x 关于 M 为(F,r)-伪凸,$r\geqslant 0$;

(ii) $\forall(x,\lambda,\mu)\in N,f$ 与 $\lambda g+\mu h$ 在 x 关于 M 分别为(F,r_1)-伪凸与(F,r_2)-拟凸,$r_1+r_2\geqslant 0$;

(iii) $\forall(x,\lambda,\mu)\in N,f$ 与 $\lambda g+\mu h$ 在 x 关于 M 分别为(F,r_1)-

拟凸与(F,r_2)-严格伪凸，$r_1+r_2\geqslant 0$.

上述的 F 如1.5.5且是给定的(下文皆如此).

证 设问题(1)与(2)非弱对偶，即有 $y\in M$，$(x,\lambda,\mu)\in N$，使 $f(y)<f(x)$，于是

$$L(y,\lambda,\mu)\leqslant f(y)<f(x)\leqslant L(x,\lambda,\mu).$$

若条件(i)满足，则以上不等式推出

$$F(y,L_x(x,\lambda,\mu))+r|y-x|=r|y-x|<0,$$

得出矛盾. 若条件(ii)满足，则分别由 $f(y)<f(x)$ 与 $\lambda g(y)+\mu h(y)\leqslant 0\leqslant\lambda g(x)+\mu h(x)$ 推出

$$F(y,f'(x))+r_1|y-x|<0$$

与
$$F(y,\lambda g'(x)+\mu h'(x))+r_2|y-x|\leqslant 0.$$

于是
$$\begin{aligned}0&=F(y,L_x(x,\lambda,\mu))\\&\leqslant F(y,f'(x))+F(y,\lambda g'(x)+\mu h'(x))\\&<-(r_1+r_2)|y-x|,\end{aligned}$$

得出矛盾. 类似地可从条件(iii)引出矛盾. □

用类似于5.7.1的方法可得出：

5.7.2 定理 以下每个条件推出问题(1)与(3)弱对偶：

(i) $\forall(x,\lambda,\mu)\in P$，$L(\cdot,\lambda,\mu)$ 在 x 关于 M 为(F,r)-伪凸，$r\geqslant 0$;

(ii) $\forall(x,\lambda,\mu)\in P$，$f$，$\lambda g$ 与 μh 在 x 关于 M 分别为(F,r_1)-凸、(F,r_2)-凸与(F,r_3)-凸，$r_1+r_2+r_3\geqslant 0$.

5.7.3 定理(逆对偶性) 设 $\bar{x}$ 与$(x^*,\lambda^*\mu^*)$分别为问题(1)与(2)的最优解且 $\bar{x}$ 为(1)的K-T点，则以下每个条件推出 $\alpha=\beta$，$\bar{x}=x^*$：

(i) $\forall(x,\lambda,\mu)\in N$，$L(\cdot,\lambda,\mu)$ 在 x 关于 M 为(F,r)-伪凸而 $L(\cdot,\lambda^*,\mu^*)$ 在 x^* 为(F,r_1)-严格伪凸，$r,r_1\geqslant 0$;

(ii) $\forall(x,\lambda,\mu)\in N$，$f$ 与 $\lambda g+\mu h$ 在 x 关于 M 分别为(F,r_1)-伪凸与(F,r_2)-拟凸，且 f 在 x^* 关于 M 为(F,r)-严格伪凸(或 $L(\cdot,\lambda^*,\mu^*)$ 在 x^* 关于 M 为(F,r)-严格伪凸)，$r\geqslant 0$，$r_1+r_2\geqslant$

$0, r + r_2 \geqslant 0$;

(iii) $\forall (x,\lambda,\mu) \in N$, f 与 $\lambda g + \mu h$ 在 x 关于 M 分别为 (F, r_1) 拟凸与 (F, r_2)- 严格伪凸, $r_1 + r_2 \geqslant 0$.

证 因条件(i),(ii),(iii) 分别蕴涵 5.7.1 中的条件(i),(ii),(iii),故 $\alpha \geqslant \beta$. 由 $\bar{x}$ 为问题(1) 的 K-T 点知有 $\bar{\lambda}, \bar{\mu}$ 使 $(\bar{x}, \bar{\lambda}, \bar{\mu}) \in N$,因此 $\alpha = f(\bar{x}) \leqslant \beta$,这推出 $\alpha = f(\bar{x}) = f(x^*) = \beta$. 设 $\bar{x} \neq x^*$. 若条件(i) 满足,则由

$$F(\bar{x}, L_x(x^*, \lambda^*, \mu^*)) + r|\bar{x} - x^*| = r|\bar{x} - x^*| \geqslant 0$$

推出 $f(\bar{x}) \geqslant L(\bar{x}, \lambda^*, \mu^*) > L(x^*, \lambda^*, \mu^*) \geqslant f(x^*)$,得出矛盾. 若条件(ii) 满足,则分别由 $f(\bar{x}) = f(x^*)$ 与 $\lambda^* g(\bar{x}) + \mu^* h(\bar{x}) \leqslant 0 \leqslant \lambda^* g(x^*) + \mu^* h(x^*)$ 推出

$$F(\bar{x}, f'(x^*)) + r|\bar{x} - x^*| < 0$$

与 $$F(\bar{x}, \lambda^* g'(x^*) + \mu^* h'(x^*)) + r_2|\bar{x} - x^*| \leqslant 0,$$

于是

$$\begin{aligned} 0 &= F(\bar{x}, L_x(x^*, \lambda^*, \mu^*)) \\ &\leqslant F(\bar{x}, f'(x^*)) + F(\bar{x}, \lambda^* g'(x^*) + \mu h'(x^*)) \\ &< -(r + r_2)|\bar{x} - x^*|, \end{aligned}$$

亦得矛盾. 当条件(iii)满足时同样可引出矛盾. □

5.7.4 定理 设 x 与 (x^*, λ^*, μ^*) 分别为问题(1) 与(3) 的最优解且 $\bar{x}$ 为(1) 的 K-T 点;5.7.2 的条件(i) 或(ii) 满足;$L(\cdot, \lambda^*, \mu^*)$ 在 x^* 关于 M 为 (F, r)- 严格伪凸, $r \geqslant 0$,则 $\bar{x} = x^*$.

下面以 $L(\cdot, \rho, \lambda, \mu) = \rho f + \lambda g + \mu h$ 取代前面的 L, 考虑问题(2)的以下变形:

$$\begin{cases} \max f(x) = \beta', \\ L_x(x, \rho, \lambda, \mu) = 0, \lambda g(x) + \mu h(x) \geqslant 0, \\ 0 \neq (\rho, \lambda, \mu) \in \mathbf{R}_+ \times Y_+^* \times Z^*. \end{cases} \tag{2'}$$

以 N' 记(2′) 的可行集. 用类似于前面的方法,可定义问题(1) 与(2′) 的弱对偶性,并得出类似于 5.7.1,5.7.3 的结果.

5.7.5 定理 设以下条件之一满足:(i) $\forall (x, \rho, \lambda, \mu) \in N'$,

$L(\cdot,\rho,\lambda,\mu)$ 在 x 关于 M 为 (F,r)-严格伪凸，$r\geqslant 0$；(ii) $\forall(x,\rho,\lambda,\mu)\in N'$，$\rho f$ 与 $\lambda g+\mu h$ 在 x 关于 M 分别为 (F,r_1) 严格伪凸与 (F,r_2)-拟凸，$r_1+r_2\geqslant 0$；(iii) $\forall(x,\rho,\lambda,\mu)\in N'$，$\rho f$ 与 $\lambda g+\mu h$ 在 x 关于 M 分别为 (F,r_1)-拟凸与 (F,r_2)-严格伪凸，$r_1+r_2\geqslant 0$. 则问题(1)与(2′)弱对偶. 若 $\bar{x}$ 与 $(x^*,\rho^*,\lambda^*,\mu^*)$ 分别为问题(1)与(2′)的最优解，且有 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in\mathbf{R}_+\times Q\times Z^*$ 使 $L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})=0$，则 $\alpha=\beta'$ 且 $\bar{x}=x^*$.

证 设条件(iii)满足(其他两种情况仿此). 若问题(1)与(2′)非弱对偶，则有 $y\in M$，$(x,\rho,\lambda,\mu)\in N'$，使 $f(y)<f(x)$. 分别由 $\rho f(y)\leqslant\rho f(x)$ 与 $\lambda g(y)+\mu h(y)\leqslant 0\leqslant\lambda g(x)+\mu h(x)$ 推出

$$F(y,\rho f'(x))+r_1|y-x|\leqslant 0$$

与
$$F(y,\lambda g'(x)+\mu h'(x))+r_2|y-x|<0,$$

于是

$$\begin{aligned}0&=F(y,L_x(x,\rho,\lambda,\mu))\\&\leqslant F(y,\rho f'(x))+F(y,\lambda g'(x)+\mu h'(x))\\&<-(r_1+r_2)|y-x|,\end{aligned}$$

得出矛盾. 因此 $\alpha\geqslant\beta'$.

其次设定理后一部分的条件满足，则 $(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})\in N'$，$\alpha=f(\bar{x})\leqslant\beta'$，因此 $\alpha=f(\bar{x})=f(x^*)=\beta'$. 若 $\bar{x}\neq x^*$，则用条件(iii)分别由 $\rho^*f(\bar{x})=\rho^*f(x^*)$ 与 $\lambda^*g(\bar{x})+\mu^*h(\bar{x})\leqslant 0\leqslant\lambda^*g(x^*)+\mu^*h(x^*)$ 推出

$$F(\bar{x},\rho^*f'(x^*))+r_1|\bar{x}-x^*|\leqslant 0$$

与
$$F(\bar{x},\lambda^*g'(x^*)+\mu^*h'(x^*))+r_2|\bar{x}-x^*|<0.$$

如同前面一样，这导致 $0<-(r_1+r_2)|\bar{x}-x^*|$！因此 $\bar{x}=x^*$. □

5.7.5 综合并推广了[171, Th. 4.1~4.3].

参考文献：[5,6,56,57,83,84,99,100,171,184].

§ 5.8 线性与二次最优化

线性最优化问题最早得到研究，且已形成很完整的理论，至少在有限维情况下是如此. 研究线性最优化问题的一些方法为非线性最优化理论所借鉴. 现在我们感兴趣的问题是：本章的一般理论涵盖了线性最优化理论的哪些特殊结果？

以下设 X, X^*, Y, Y^* 中分别由 X_+, X_+^*, Y_+, Y_+^* 导入的序皆记为 $\leqslant$. 给定 $f \in X^*$, $A \in L(X, Y)$, $a \in Y$, 考虑**线性最优化问题**

$$\min f(x) = \alpha,\ Ax \leqslant a,\ x \geqslant 0. \tag{1}$$

首先考虑对问题(1)应用 5.2.3. 依 5.2.3 的记号有 $g(x) = Ax - a$, $D = X_+$, $L(x,\lambda) = f(x) + \langle \lambda, Ax - a\rangle$. 因

$$\inf L(D,\lambda) = \inf \langle A^*\lambda + f, X_+\rangle - \lambda(a) = \begin{cases} -\lambda(a), & A^*\lambda + f \geqslant 0, \\ -\infty, & \text{否则}, \end{cases}$$

故问题(1)的 Lagrange 对偶是

$$\max [-\lambda(a)] = \beta,\ A^*\lambda \geqslant -f,\ \lambda \geqslant 0. \tag{1*}$$

注意问题(1*)是与问题(1)同类型的线性最优化问题. 分别以 M 与 N 记问题(1)与(1*)的可行集，则 $M \neq \varnothing \Leftrightarrow \alpha < \infty$, $N \neq \varnothing \Leftrightarrow \beta > -\infty$. 条件

$$f(\bar{x}) = \min L(D, \bar{\lambda}),\ \langle \bar{\lambda}, g(\bar{x})\rangle = 0,\ (\bar{x}, \bar{\lambda}) \in M \times Y_+^*$$

可表成如下对称形式：

$$\begin{cases} \langle \bar{\lambda}, A\bar{x} - a\rangle = \langle A^*\bar{\lambda} + f, \bar{x}\rangle = 0, \\ \bar{x} \geqslant 0, \bar{\lambda} \geqslant 0, A\bar{x} \leqslant a, A^*\bar{\lambda} \geqslant -f. \end{cases} \tag{2}$$

式(2)就是线性规划中有名的**松弛互补条件**.

另一方面，若作扰动函数

$$\varphi(x, y) = \begin{cases} f(x), & Ax + y \leqslant a, x \geqslant 0, \\ \infty, & \text{否则}, \end{cases}$$

则依§5.4之式(1)不难算得

$$S(y)=\inf\{f(x):Ax\leqslant a-y,x\geqslant 0\}\overset{\text{def}}{=}\alpha(a-y);$$

$$\varphi^*(u,\lambda)=\begin{cases}\lambda(a), & A^*\lambda\geqslant u-f,\lambda\geqslant 0,\\ \infty, & \text{否则};\end{cases}$$

$$S_*(u)=\inf\{\lambda(a):A^*\lambda\geqslant u-f,\lambda\geqslant 0\}\overset{\text{def}}{=}-\beta(f-u).$$

由此得出 $\partial S(0)=-\partial\alpha(a)$, $\partial S_*(0)=-\partial(-\beta)(f)$.

结合以上结论与5.2.3,5.4.2得:

5.8.1 定理 对问题(1)与(1*)以下结论成立:

(i) $(\bar{x},\bar{\lambda})$是$L(x,\lambda)$在$X_+\times Y_+^*$上的鞍点$\Leftrightarrow$ $\bar{x}$与$\bar{\lambda}$分别为问题(1)与(1*)的解且$\alpha=\beta$(必定$\alpha\neq\pm\infty$)$\Leftrightarrow$条件式(2)满足.

(ii) 设$M,N\neq\varnothing$. 则问题(1)稳定$\Leftrightarrow$问题(1*)有解且$\alpha=\beta$(此时问题(1*)的解集为$-\partial\alpha(a)$);若X,Y自反,则问题(1*)稳定$\Leftrightarrow$问题(1)有解且$\alpha=\beta$(此时问题(1)的解集为$-\partial(-\beta)(f)$).

(iii) 若$N\neq\varnothing$, $\exists\hat{x}\geqslant 0$: $A\hat{x}\ll a$,则问题(1)稳定,从而问题(1*)有解且$\alpha=\beta$;若$M\neq\varnothing$, $\exists\hat{\lambda}\geqslant 0$: $A^*\hat{\lambda}\gg -f$,则问题(1*)稳定,从而当X,Y自反时问题(1)有解且$\alpha=\beta$.

5.8.1虽不失为一个有价值的结果,但并未达到线性规划中相应结果的完满程度. 更强的有限维结果如下.

5.8.2 定理 设$X=\mathbf{R}^n$, $X_+=\mathbf{R}_+^n$, $Y=\mathbf{R}^m$, $Y_+=\mathbf{R}_+^m$. 若M, $N\neq\varnothing$,则问题(1)与(1*)皆有解、稳定且$\alpha=\beta$,若$M=\varnothing$或$N=\varnothing$,则(1)与(1*)皆无解.

以上结果的证明基于某些精细的线性代数方法,这些方法无法推广于无限维情况.

现在转而考虑二次最优化问题

$$\min\left[\frac{1}{2}\langle Qx,x\rangle+w(x)+c\right]=\alpha,\ Ax\leqslant a, \tag{3}$$

其中$Q=Q^*\in L(X,X^*)$, $w\in X^*$, $c\in\mathbf{R}$,假定X,Y自反. 令

$$f(x)=\frac{1}{2}\langle Qx,x\rangle+w(x)+c,$$

$$L(x,\lambda) = f(x) + \langle \lambda, Ax - a \rangle.$$

设 Q 强单调(1.3.2),则 Q 可逆(1.3.7).因

$$L_x(x,\lambda) = f'(x) + A^*\lambda = Qx + w + A^*\lambda$$

对 x 严格单调,$L_x(x,\lambda) = 0$ 有唯一解 $x = -Q^{-1}(w + A^*\lambda)$,故由 1.5.3,1.6.3 有

$$\begin{aligned}
\min_{x\in X} L(x,\lambda) &= L(-Q^{-1}(w + A^*\lambda),\lambda) \\
&= \frac{1}{2}\langle w + A^*\lambda, Q^{-1}(w + A^*\lambda)\rangle \\
&\quad - \langle w + A^*\lambda, Q^{-1}(w + A^*\lambda)\rangle + c - \lambda(a) \\
&= -\frac{1}{2}\langle \lambda, P\lambda\rangle - \langle \lambda, b\rangle - \tilde{c},
\end{aligned}$$

其中 $P = AQ^{-1}A^* = P^* \in L(Y^*,Y)$,$b = AQ^{-1}w + a$,$\tilde{c} = \frac{1}{2}\langle w, Q^{-1}w\rangle - c$. 于是问题(3)的 Lagrange 对偶为

$$\max\left[-\frac{1}{2}\langle \lambda, P\lambda\rangle - \langle \lambda, b\rangle - \tilde{c}\right] = \beta,\ \lambda \geqslant 0. \qquad (3^*)$$

注意问题(3^*)的目标函数与问题(3)一样是"二次函数",但约束较简单. 以 M 记(3)的可行集,则 $M \neq \varnothing \Leftrightarrow \alpha$ 有限. 另一方面,恒有 $\beta > -\infty$. 条件"$L_x(\bar{x},\bar{\lambda}) = 0, \bar{x} \in M, \bar{\lambda} \geqslant 0, \bar{\lambda}g(\bar{x}) = 0$"可表成

$$\begin{cases} Q\bar{x} + A^*\bar{\lambda} + w = 0, \langle \bar{\lambda}, A\bar{x} - a\rangle = 0, \\ A\bar{x} \leqslant a, \bar{\lambda} \geqslant 0. \end{cases} \qquad (4)$$

若作扰动函数

$$\varphi(x,y) = \begin{cases} f(x), & Ax + y \leqslant a, \\ \infty, & 否则, \end{cases}$$

则 $$S(y) = \inf\{f(x) : Ax \leqslant a - y\} \overset{\text{def}}{=} \alpha(a - y);$$

$$\begin{aligned}
\varphi^*(u,\lambda) &= \sup_{Ax+y\leqslant a}[u(x) + \lambda(y) - f(x)] \\
&= \sup_{x\in X, z\geqslant 0}[u(x) + \lambda(a - Ax - z) - f(x)]
\end{aligned}$$

$$= \sup_{x \in X}\left[\langle u - w - A^{*}\lambda, x\rangle - \frac{1}{2}\langle Qx, x\rangle\right]$$

$$- \inf_{z \geqslant 0} \lambda(z) - c + \lambda(a)$$

$$= \begin{cases} \frac{1}{2}\langle u - w - A^{*}\lambda, Q^{-1}(u - w - A^{*}\lambda)\rangle \\ \quad + \lambda(a) - c, & \lambda \geqslant 0, \\ \infty, & \text{否则} \end{cases}$$

(参考 5.3.2 之 3°),于是

$$S_{*}(u) = \inf\left\{\frac{1}{2}\langle u - w - A^{*}\lambda, Q^{-1}(u - w - A^{*}\lambda)\rangle\right.$$

$$\left. + \lambda(a) - c : \lambda \geqslant 0\right\} \stackrel{\text{def}}{=} -\beta(w - u).$$

由此得出 $\partial S(0) = -\partial\alpha(a)$,$\partial S_{*}(0) = -\partial(-\beta)(w)$. 结合以上结论与 5.2.3,5.4.2 得出:

5.8.3 定理 设 X,Y 为自反 B-空间,$Q = Q^{*} \in L(X, X^{*})$ 强单调,$w \in X^{*}$,则以下结论成立:

(i) $\bar{x}$ 与 $\bar{\lambda}$ 分别为问题(3)与(3*)的解且 $\alpha = \beta \Leftrightarrow$ 条件式(4)成立 $\Leftrightarrow (\bar{x}, \bar{\lambda})$ 是 $L(x, \lambda)$ 在 $X \times Y_{+}^{*}$ 上的鞍点.

(ii) 若 $M \neq \varnothing$,则问题(3)有解且 $\alpha = \beta$;若 $\exists \hat{x} \in X : A\hat{x} \ll a$,则问题(3)与(3*)皆有解且 $\alpha = \beta$,从而 $L(x, \lambda)$ 在 $X \times Y_{+}^{*}$ 上有鞍点.

(iii) 若 $\exists \hat{x} \in X : A\hat{x} \ll a$,则问题(3)稳定,问题(3*)有非空解集 $-\partial\alpha(a)$;若 $M \neq \varnothing$,则问题(3*)稳定,问题(3)有非空解集 $-\partial(-\beta)(w)$.

第六章　向量最优化

本章考虑如下的向量最优化问题(VOP)：

(P)：　$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D$,

其中约束(因而可行集 M)与本书前面考察的“标量最优化问题”并无不同，所不同的只是目标函数. 在问题(P)中 $f: X \to W$，其中 W 由闭凸锥 W_+ 导入序 $\leqslant$. 不失一般性，假定 W_+ 是序锥，即 $W_+ \cap W_- = \{0\}$.

本章沿用前几章的记号：$M = D \cap G \cap H, G = g^{-1}(Y_-), H = h^{-1}(0), K_g = Y_- + \mathbf{R}g(\bar{x}), \bar{x} \in M; Q = -K_g^* = Y_+^* \cap \{g(\bar{x})\}^\perp; T = f'(\bar{x}), A = g'(\bar{x}), B = h'(\bar{x})$，只要它们存在；$L(\cdot, \rho, \lambda, \mu) = \rho f + \lambda g + \mu h$. 若 $W = \mathbf{R}^n$，则约定 $W_+ = \mathbf{R}_+^n, f = (f_i)$. 若 $Y = \mathbf{R}^m$，则约定 $Y_+ = \mathbf{R}_+^m, g = (g_i), I = \{i : g_i(\bar{x}) = 0\}$，从而 $Q = \mathbf{R}_+^I$.

§6.1　向量极值

设 $A \subset W$. 令

$$\min A = \{a \in A : A \cap (a - W_+) = \{a\}\} \tag{1}$$
$$(= \{a \in A : (A - a) \cap W_-) = \{0\}\});$$

$$\text{wmin } A = \{a \in A : A \cap (a - W_+^\circ) = \varnothing\} \tag{2}$$
$$(= \{a \in A : (A - a) \cap W_-^\circ) = \varnothing\});$$

$$\text{pmin } A = \{a \in A : \overline{A + W_+} \cap (a - W_+) = \{a\}\} \tag{3}$$
$$(= \{a \in A : \overline{A - a + W_+} \cap W_-) = \{0\}\}).$$

分别称 $\min A$, $\operatorname{wmin} A$ 与 $\operatorname{pmin} A$ 中的点为 A 的极小点、弱极小点与真极小点. 对于 $a \in A$, 由式(1), (2) 看出 $a \in \min A \Leftrightarrow \forall b \in A: b \not< a$, $a \in \operatorname{wmin} A \Leftrightarrow \forall b \in A: b \not\ll a$. 约定

$$\inf A = \min \overline{A},\ \operatorname{winf} A = \operatorname{wmin} \overline{A};$$

$$\max A = -\min(-A),\ \sup A = \max \overline{A}.$$

$\operatorname{wmax} A$, $\operatorname{pmax} A$, $\operatorname{wsup} A$ 等仿此.

今将 $\min A$, $\operatorname{wmin} A$ 等的性质汇集如下.

6.1.1 命题 对 $A, B \subset W$, $b \in W$ 以下结论成立:

(i) $\operatorname{pmin} A \subset \min A = \min(A + W_+) \subset \operatorname{wmin} A = \operatorname{wmin}(A + W_+)$.

(ii) $\operatorname{pmin} A \subset A \cap \inf A \subset \min A$; $\operatorname{wmin} A \subset \operatorname{winf} A$; $A \cap \operatorname{winf} A \subset \operatorname{wmin} A$.

(iii) $\min(A + B) \subset \min A + \min B$; $\min(A + b) = \min A + b$; $\operatorname{wmin} A$ 等仿此.

(iv) $A + W_+$ 闭 $\Rightarrow \operatorname{pmin} A = \min A = \inf A$.

(v) $\min A \subset \partial A$; $\operatorname{wmin} A \cup \inf A \cup \operatorname{winf} A \subset \partial \overline{A}$.

证明都是直接的.

6.1.2 例 取 $W = \mathbf{R}^2$, $W_+ = \mathbf{R}^2_+$.

1° 设 $A = \{0\} \cup \{x \in \mathbf{R}^2 : x_1 x_2 = -1, x_2 < 0\}$, $B = (-\infty, 0] \times \{0\}$, 则 $\operatorname{pmin} A = \varnothing$, $\min A = \{0\}$; $\min B = \varnothing$; $\operatorname{wmin} B = B$. 可见 6.1.1(i) 中的包含不必是等式.

2° 设 $A = \mathring{\mathbf{R}}^2_+$, $B = ((-\infty, 0) \times (0, \infty)) \cup \{0\}$, 则 $\min A = \varnothing$, $\inf A = \{0\}$, $\min B = \{0\}$, $\inf B = \varnothing$. 可见 $\min A$ 与 $\inf A$ 之间不必有包含关系.

6.1.3 定理 设 $a \in A \subset W$.

(i) 若 $\exists \rho \in W_+^* \setminus \{0\}: \rho(a) = \min \rho(A)$, 则 $a \in \operatorname{wmin} A$; 当 $W_+^\circ \neq \varnothing$ 且 $A + W_+$ 凸时其逆亦真.

(ii) 若 $\exists \rho \in \mathring{W}_+^*: \rho(a) = \min \rho(A)$, 则 $a \in \operatorname{pmin} A$; 当 $\mathring{W}_+^* \neq$

$\varnothing$, W 自反且 $A+W_+$ 凸时，其逆亦真.

(iii) 若 $\mathring{W}_+^* \neq \varnothing$，$W$ 自反且 $A+W_+$ 闭凸，则 $a \in \min A \Leftrightarrow a \in \operatorname{pmin} A \Leftrightarrow \exists \rho \in \mathring{W}_+^*: \rho(a) = \min \rho(A)$.

证 (i) 若 $a \overline{\in} \operatorname{wmin} A$，则 $\exists b \in A: b \ll a$. 于是 $\forall \rho \in W_+^* \backslash \{0\}$，有 $\rho(b) < \rho(a)$(1.4.4)，因而 $\rho(a) \neq \min \rho(A)$. 反之，若 $a \in \operatorname{wmin} A$，则 $(A+W_+) \cap (a - \mathring{W}_+) = \varnothing$. 于是当 $\mathring{W}_+ \neq \varnothing$ 且 $A+W_+$ 凸时 $\exists \rho \in W^* \backslash \{0\}: \langle \rho, a - \mathring{W}_+ \rangle \leqslant \langle \rho, A + W_+ \rangle$ (1.1.2)，这推出 $\rho \in W_+^* \backslash \{0\}, \rho(a) = \min \rho(A)$.

(ii) 设 $\rho \in \mathring{W}_+^*$ 使 $\rho(a) = \min \rho(A)$. $\forall b \in \overline{A+W_+} \cap (a - W_+)$，有 $b \leqslant a$，且 $\exists a_n \in A, x_n \geqslant a_n$，使 $x_n \to b$. 于是

$$
\begin{aligned}
\rho(b) &\leqslant \rho(a) \leqslant \limsup_n \rho(a_n) \\
&\leqslant \lim_n \rho(x_n) = \rho(b),
\end{aligned}
$$

这推出 $a=b$. 因此 $a \in \operatorname{pmin} A$. 反之，若 $\mathring{W}_+^* \neq \varnothing$，$\forall \rho \in \mathring{W}_+^*: \rho(a) > \inf \rho(A)$，则 $\mathring{W}_+^* \cap (A-a)^* = \varnothing$. 于是当 W 自反时有 $x \in W \backslash \{0\}$，使 $\langle W_+^*, x \rangle \leqslant \langle (A-a)^*, x \rangle$. 这推出 $x \in W_-$；再加上 $A+W_+$ 凸有(1.4.3)

$$x \in (A-a)^{**} \subset (A+W_+-a)^{**} = \overline{A+W_+-a},$$

于是 $a \neq a + x \in \overline{A+W_+} \cap (a - W_+)$，可见 $a \overline{\in} \operatorname{pmin} A$.

(iii) 只需注意 $A+W_+$ 闭 $\Rightarrow \min A = \operatorname{pmin} A$，其余由(ii)推出. □

现在将已给的概念与结论转移到**向量最优化**问题(VOP). 设 $f: X \to W$，考虑 VOP

$$\min f(x), \ x \in M. \tag{4}$$

6.1.4 定义 设 $\bar{x} \in M$. 若 $f(\bar{x}) \in \min f(M)$(即 $\forall x \in M$: $f(x) \not< f(\bar{x})$)，则称 $\bar{x}$ 为问题(4)的**有效解**(或 **Pareto 最优解**). 若 $f(\bar{x}) \in \operatorname{wmin} f(M)$(即 $\forall x \in M: f(x) \not\ll f(\bar{x})$)，则称 $\bar{x}$ 为问题

(4) 的**弱有效解**(或 Slater 最优解). 若 $f(\bar{x}) \in \mathrm{pmin}\, f(M)$,则称 $\bar{x}$ 为问题(4) 的**真有效解**. 若 $\forall x \in M\setminus\{\bar{x}\}: f(x) \not\leqslant f(\bar{x})$,则称 $\bar{x}$ 为问题(4) 的**严格有效解**.

局部[弱]有效解等概念的意义是自明的.

结合 6.1.1 与 6.1.4 直接看出:严格有效解与真有效解是有效解,而有效解必为弱有效解. 若 $W = \mathbf{R}$,则真有效解、有效解、弱有效解都相当于最优解.

研究 VOP(4)的一个重要方法是将它转化为一个标量最优化问题(更一般的处理参看 § 6.4)

$$\min \rho f(x),\ x \in M, \tag{4$_\rho$}$$

其中 $\rho \in W_+^*$. 取 $A = f(M)$ 从 6.1.3 直接得到:

6.1.5 定理 设 $\bar{x} \in M$.

(i) 若 $\exists \rho \in W_+^* \setminus \{0\}$,使 $\bar{x}$ 是问题$(4)_\rho$ 的最优解,则 $\bar{x}$ 是 VOP(4) 的弱有效解;当 $\mathring{W}_+^* \neq \varnothing$ 且 $f|M$ 为 W_+ - 类凸时其逆亦真.

(ii) 若 $\exists \rho \in \mathring{W}_+^*$,使 $\bar{x}$ 是问题$(4)_\rho$ 的最优解,则 $\bar{x}$ 是 VOP(4) 的真有效解;当 W 自反,$\mathring{W}_+^* \neq \varnothing$ 且 $f|M$ 为 W_+ - 类凸时,其逆亦真.

(iii) 若 W 自反,$\mathring{W}_+^* \neq \varnothing$,$f|M$ 为 W_+ - 类凸且 $f(M) + W_+$ 闭,则 $\bar{x}$ 是 VOP(4) 的有效解 $\Leftrightarrow$ $\bar{x}$ 是 VOP(4) 的真有效解 $\Leftrightarrow \exists \rho \in \mathring{W}_+^*: \bar{x}$ 是问题$(4)_\rho$ 的最优解.

(iv) 若 $\exists \rho \in W_+^*$,使 $\bar{x}$ 是问题$(4)_\rho$ 的严格最优解,则 $\bar{x}$ 是 VOP(4) 的严格有效解.

显然,将 6.1.5 中的"解"改成"局部解"之后,结论仍然成立. 其次,由 3.2.7 得出,6.1.5(iii) 中的条件"$f|M$ 为 W_+ - 类凸"可减弱为"$f|M$ 为 W_+ - 次类凸".

以下结果可看作 1.6.3 的向量形式.

6.1.6 定理　设 $\bar{x} \in M$.

(i) 若 X 自反，M 有界凸闭，$\mathring{W}_+^* \neq \varnothing$，$f$ 为 $*$ 拟凸且为 $*$ lsc，则 VOP(4) 有真有效解.

(ii) 若 $f|M$ 为 W_+ - 凸，则 VOP(4) 的局部[弱]有效解是[弱]有效解.

(iii) 若 $f|M$ 为 W_+ - 凸，对任给 $x \in M$，有 $\mathrm{D}_+ f(\bar{x}, x - \bar{x}) \not< 0$ [$\mathrm{D}_+ f(\bar{x}, x - \bar{x}) \not\ll 0$]，则 $\bar{x}$ 是 VOP(4) 的[弱]有效解.

(iv) 若 f 在 $\bar{x}$ 关于 M 为 $*$ 伪凸，$T = f'(\bar{x})$ 存在，$\mathring{W}_+^* \cap N(T^*) \neq \varnothing$[$(W_+^* \setminus \{0\}) \cap N(T^*) \neq \varnothing$]，则 $\bar{x}$ 是 VOP(4) 的真有效解[弱有效解].

证　(i) 由 1.6.3(i) 与 6.1.5(ii) 得出.

(ii) 的证明类似于 1.6.3(iv).

(iii) 若 $\bar{x}$ 不是 VOP(4) 的有效解，则 $\exists x \in M$：$f(x) < f(\bar{x})$. 这结合 $f|M$ 凸得出：

$$
\begin{aligned}
\mathrm{D}_+ f(\bar{x}, x - \bar{x}) &= \lim_{t \downarrow 0} \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t} \\
&\leqslant \lim_{t \downarrow 0} \frac{t' f(\bar{x}) + t f(x) - f(\bar{x})}{t} \\
&= f(x) - f(\bar{x}) < 0.
\end{aligned}
$$

类似地，$\bar{x}$ 不是 VOP(4) 的弱有效解 $\Rightarrow \mathrm{D}_+ f(\bar{x}, x - \bar{x}) \ll 0$.

(iv) 取 $\rho \in \mathring{W}_+^* \cap N(T^*)$. 由 $f_\rho \overset{\mathrm{def}}{=} \rho f$ 伪凸与 $f'_\rho(\bar{x}) = T^* \rho = 0$ 得出 $\rho f(\bar{x}) = \min \rho f(M)$(1.6.3(vi))，于是由 6.1.5(ii) 知 $\bar{x}$ 是 VOP(4) 的真有效解. 关于弱有效解的结论可类似证明. □

6.1.7 定理　设 W 自反，W_+ 是正规的序锥，$\mathring{W}_+^* \neq \varnothing$；$M$ 为紧集，f 在 M 上关于 W_+ 为 lsc，则 $\forall x \in M$，存在问题 VOP(4) 的真有效解 $\bar{x}$，使得 $f(x) \geqslant f(\bar{x})$.

证　令 $A = f(M)$，只要证 $A \subset \mathrm{pmin}\, A + W_+$.

1° 证存在有界集 $V \subset W$，使 $A \subset V + W_+$. 否则，$\forall n \geqslant 1, \exists x_n \in M$：$f(x_n) \overline{\in} B_n(0) + W_+$. 因 M 紧，不妨设 $x_n \to x \in M$. 因 f 在 x 为 lsc，故当 n 充分大时

$$f(x_n) \in B_1(f(x)) + W_+ \subset B_n(0) + W_+,$$

得出矛盾.

2° 证 $A + W_+$ 弱闭. 设 $x_n \in M, y_n \in W_+, f(x_n) + y_n \rightharpoonup z$. 可设 $x_n \to x \in M$. 由 f 在 x 为 lsc，可设 $f(x_n) \geqslant a_n \to f(x)$，从而

$$f(x_n) + y_n - a_n \rightharpoonup z - f(x) \in W_+,$$

于是 $z = f(x) + (z - f(x)) \in A + W_+$.

3° 任取 $a \in A$，证存在 $b \leqslant a$：$b \in \operatorname{pmin} A$. 令

$$B = (A + W_+) \cap (a - W_+).$$

由 2°，B 弱闭；由 W_+ 正规得出 B 有界([93,Th. 2. 1. 4])，因此弱紧(1. 1. 6). 取 $\rho \in \mathring{W}_+^*$，则有 $b \in B$：$\rho(b) = \min \rho(B)$(1. 6. 2)，于是 $b \in \operatorname{pmin} B$(6. 1. 3)，$b \leqslant a$. 设 $b = a' + y = a - z, a' \in A, y, z \in W_+$. 若 $y \neq 0$，则 $b > a' = a - (y + z) \in B$，这与 $b \in \operatorname{pmin} B$ 矛盾. 因此 $b = a' \in A$. 若 $b \overline{\in} \operatorname{pmin} A$，则由式(3)有 $(A + W_+) \cap (b - W_+) \neq \{b\}$，于是有 $b \neq c = a' + y = b - z, a' \in A, y, z \in W_+$. 这推出 $b > b - z \in (A + W_+) \cap (a - W_+ - z) \subset B$，这与 $b \in \operatorname{pmin} B$ 矛盾. 因此 $b \in \operatorname{pmin} A$. □

注 1° 若 $W = \mathbf{R}$，$\bar{x}$ 是 VOP(4) 的最优解，则无需任何附加条件就有 $f(x) \geqslant f(\bar{x})(\forall x \in M)$. 6. 1. 7 显示出向量极值问题的复杂性.

2° 6. 1. 7 中条件“M 紧，f 在 M 上为 lsc”可改为“M 有界闭，f 在 M 上为 wlsc”. f 在 x 为 wlsc 意味着：若 $x_n \rightharpoonup x$，则 $\exists y_n \to f(x)$：$f(x_n) \geqslant y_n$.

3° 在 6. 1. 7 的条件下，实际上我们已证明了

$$A + W_+ = \min A + W_+.$$

6. 1. 8 引理 设 $B \subset W, A = \overline{B + W_+}$，则 ∂A 中任两点 a, b

不依 $\ll$ 相关.

证 设 $a,b \in \partial A, a \ll b$,则

$$b \in a + W_+^\circ \subset A + W_+^\circ \subset (A + W_+)^\circ$$
$$= (\overline{B + W_+} + W_+)^\circ \subset (\overline{B + W_+ + W_+})^\circ$$
$$\subset (\overline{B + W_+})^\circ = A^\circ,$$

这与 $b \in \partial A$ 相矛盾. □

参考文献:[27,45,53,68,79,93,143,210,211,241].

§6.2 最优性条件

本节展开平行于 §4.2~§4.5 的内容,凡属例行的推广皆省略证明.

考虑如下 VOP:

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D. \tag{1}$$

设 $\bar{x} \in M = D \cap G \cap H$, f,g,h 皆在 $\bar{x}$ 可微, $L(\cdot,\rho,\lambda,\mu) = \rho f + \lambda g + \mu h$. 记号 T,A,B,K_g,Q 等皆依旧. 注意

$$L_x(\bar{x},\rho,\lambda,\mu) = T^*\rho + A^*\lambda + B^*\mu.$$

推广 §4.2(2),定义

$$F_f(x) = \{z: \text{当 } 0 < t \downarrow 0, w \to z \text{ 时 } f(\bar{x} + tw) \ll f(\bar{x})\}. \tag{2}$$

$F_f(\bar{x})$ 是一开锥;当 f 为 W_+ - 凸时 $F_f(\bar{x})$ 凸. 若 $W = \mathbf{R}^n$,则

$$F_f(\bar{x}) = \bigcap_{l=1}^{n} F_{f_l}(\bar{x}).$$

6.2.1 引理 若 $\bar{x}$ 是 VOP(1) 的局部弱有效解,则

$$F_f(\bar{x}) \cap K_M(\bar{x}) = \varnothing.$$

6.2.2 引理 $T^{-1}W_-^\circ \subset F_f(\bar{x})$; 若 $T^{-1}W_-^\circ \neq \varnothing$, 则

$$T^{-1}W_-^\circ = F_f(\bar{x}).$$

证 只需证后者. 任给 $z \in F_f(\bar{x})$, 取 $\hat{x} \in T^{-1}W_-^\circ$, 可设 $z -$

$\hat{x} \approx z$. 若 $z \bar{\in} T^{-1}W_{-}^{\circ}$，即 $Tz \bar{\in} W_{-}^{\circ}$，则由 1.1.2 有 $0 \neq \rho \in W^{*}$：$\langle \rho, Tz \rangle \geqslant \langle \rho, W_{-} \rangle$，因此 $\langle \rho, Tz \rangle \geqslant 0, \rho \in W_{+}^{*}$. 于是当 $0 < t \downarrow 0$ 时有

$$0 \leqslant \langle \rho,\ tTz + f(\bar{x}) - f(\bar{x} + t(z - \hat{x})) \rangle$$
$$= \langle \rho,\ tT\hat{x} + o(t) \rangle < 0,$$

得出矛盾. 故 $F_f(\bar{x}) \subset T^{-1}W_{-}^{\circ}$，因而 $T^{-1}W_{-}^{\circ} = F_f(\bar{x})$. □

6.2.3 Fritz John 定理 设 $W_{+}^{\circ} \neq \varnothing \neq Y_{+}^{\circ}$，$K \subset F_D(\bar{x})$ 是一非空开凸锥；h 在 $\bar{x}$ 邻近为 C^1 映射，$R(B)$ 为闭集. 若 $\bar{x}$ 是(1) 的局部弱有效解，则有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in W_{+}^{*} \times Q \times Z^{*}$，使 $L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) \in K^{*}$；特别，当 D 凸且 $D^{\circ} \neq \varnothing$ 时 $L_x(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) \in (D - \bar{x})^{*}$. 若 $Y = \mathbf{R}^m$，则取 $Q = \mathbf{R}_{+}^{l}$.

证明略同于 4.2.3，只是此处取 $K_0 = T^{-1}W_{-}^{\circ}$. 可设 $K_0 \neq \varnothing$(否则由 3.3.3 有 $0 \neq \bar{\rho} \in W_{+}^{*}: T^{*}\bar{\rho} = 0$，取 $\bar{\lambda} = 0, \bar{\mu} = 0$ 好了)，于是 $K_0^{*} = -T^{*}W_{+}^{*}$ (4.1.2).

取 $Z = \{0\}$，6.2.3 自然用到问题

$$\min f(x),\ g(x) \leqslant 0,\ x \in D. \tag{3}$$

6.2.4 定理 设 $W_{+}^{\circ} \neq \varnothing$，$K \subset F_D(\bar{x})$ 是一非空开凸锥；$\bar{x}$ 是(3) 的局部弱有效解. 若 $Y_{+}^{\circ} \neq \varnothing$，或 $g \in C^1$ 且 $R(A) + K_g = Y$，则有 $0 \neq (\bar{\rho}, \bar{\lambda}) \in W_{+}^{*} \times Q$，使 $T^{*}\bar{\rho} + A^{*}\bar{\lambda} \in K^{*}$.

当 $Y_{+}^{\circ} \neq \varnothing$ 时结论直接由 6.2.3 得出；余下部分的证明略同于 4.2.6.

6.2.5 Dubovickii-Miljutin 定理 设 $M_j \subset X (1 \leqslant j \leqslant m)$，$\bar{x} \in M = \bigcap M_j$，$K_0 = F_f(x)$，$K_j = F(M_j, \bar{x}) (1 \leqslant j < m)$，$K_m = K(M_m, \bar{x})$，$K_j (0 \leqslant j \leqslant m)$ 非空凸.

(i) 若 $\bar{x}$ 是问题

$$\min f(x),\ x \in M_j,\quad j = 1, 2, \cdots, m \tag{4}$$

的局部弱有效解，则 $\{K_j^{*}: 0 \leqslant j \leqslant m\}$ 为 SOS(参考 §4.2).

(ii) 若 M_j 皆凸，$M_1^{\circ} \cap \cdots \cap M_{m-1}^{\circ} \cap M_m \neq \varnothing$，$f$ 为 usc(参考

3.4.2)且 * 严格拟凸,$\{K_j^*:0\leqslant j\leqslant m\}$ 为 SOS,则 $\bar{x}$ 是问题(4)的弱有效解.

证明如同 4.2.7. 若 $W=\mathbf{R}^n, K_{m+i}\overset{\text{def}}{=}F_{f_i}(\bar{x})$ 非空凸,则 6.2.5 中的必要条件可表为 $\{K_j^*:0\leqslant j\leqslant m+n\}$ 为 SOS. 进一步的研究可参看[126,127,130,131,154].

以下设 $D=X, \bar{x}\in M=G\cap H$.

6.2.6 定义 若存在 $(\bar{\rho},\bar{\lambda},\bar{\mu})\in(W_+^*\setminus\{0\})\times Q\times Z^*$,使 $L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})=0$,则称 $\bar{x}$ 为 K-T 点,称 $\bar{\rho},\bar{\lambda},\bar{\mu}$ 为 K-T 乘子;若上述的 $\bar{\rho}\in\mathring{W}_+^*$,则称 $\bar{x}$ 为强 K-T 点.

由 6.2.3 直接推出(参照 4.3.2):

6.2.7 定理 设 $W_+^\circ\neq\varnothing\neq Y_+^\circ$, h 在 $\bar{x}$ 邻近为 C^1 映射且 $R(B)=Z$. 若 $\bar{x}$ 是(1)的局部弱有效解,则以下每个条件推出 $\bar{x}$ 为 K-T 点:

(i) $(A,B):X\to Y\times Z$ 为满射;

(ii) $N(B)\cap A^{-1}(Y_-^\circ+\mathbf{R}g(\bar{x}))\neq\varnothing$;

(iii) g 为 * 伪凸, $h(x)=Bx-b$,且存在 $\hat{x}\in X$,使 $g(\hat{x})\ll 0, h(\hat{x})=0$.

下面作类似于引出 4.3.4 的分析. 首先注意到, $\bar{x}$ 是(1)的 K-T 点相当于

$$T^*(W_-^*\setminus\{0\})\cap(A^*Q+R(B^*))\neq\varnothing. \tag{5}$$

设 $K\subset X$ 是一非空凸锥, $W_+^\circ\neq\varnothing$. 今指明(6)蕴涵于条件

$$K\cap T^{-1}W_-^\circ=\varnothing,\ -K^*\subset A^*Q+R(B^*). \tag{6}$$

事实上,若 $T^{-1}W_-^\circ=\varnothing$,则从 6.2.3 之证已可见 $\bar{x}$ 为 K-T 点. 若 $T^{-1}W_-^\circ\neq\varnothing$,则 $(T^{-1}W_-^\circ)^*=T^*W_-^*$ 与 K^* 为 SOS(4.2.2),于是有 $0\neq u\in(-K^*)\cap T^*W^*$,由此看出(6) $\Rightarrow$ (5). 基于此,有与 4.3.4 相当的以下结果.

6.2.8 定理 设 $W_+^\circ\neq\varnothing$, $\bar{x}$ 是(1)的局部弱有效解. 则以下每个条件推出 $\bar{x}$ 为 K-T 点:

(i) $-K_M(\bar{x})=N(B)\cap A^{-1}(Y\cap Q^*)$ 且 $A^*Q+R(B^*)$ 弱*闭；

(ii) $g,h\in C^1,(A,B):X\to Y\times Z$ 是满射，K_g 闭且 $A^*Q+R(B^*)$ 弱*闭；

(iii) $K_M(\bar{x})=N(B)\cap A^{-1}K_g$，$K_g$ 与 $R(B)$ 闭，$R(A)+K_g=Y$ 且 $N(B)+A^{-1}K_g=X$.

证明略同于 4.3.4，只要取 $K=K_M(\bar{x})$ 验证(6).

将 6.2.7，6.2.8 用到问题(3)($D=X$) 得出：

6.2.9 推论　设 $W_+^{\circ}\neq\varnothing$，$\bar{x}$ 是式(3) 的局部弱有效解. 则以下每个条件推出 $\bar{x}$ 为 K-T 点：

(i) $Y_+^{\circ}\neq\varnothing,R(A)=Y$；

(ii) $A^{-1}(Y_-^{\circ}+\mathbf{R}g(\bar{x}))\neq\varnothing$；

(iii) g 为 * 伪凸，$\exists\hat{x}\in X:g(\hat{x})\ll 0$；

(iv) $A^{-1}(Y\cap Q^*)\cap T^{-1}W_+^*=\varnothing$ 且 A^*Q 弱*闭；

(v) $g\in C^1,R(A)+K_g=Y$.

下面考虑运用择一定理. 令 $\tilde{f}(x)=f(x)-f(\bar{x})$，则 $\bar{x}$ 是 (1)，(3) 的弱有效解分别意味着以下问题无解：

$$\exists x\in D:\tilde{f}(x)\ll 0,g(x)\leqslant 0,h(x)=0;\tag{7}$$

$$\exists x\in D:\tilde{f}(x)\ll 0,g(x)\leqslant 0.\tag{8}$$

问题(7)可改写成：

$$\exists x\in D\cap H:\tilde{f}(x)\ll 0,g(x)\leqslant 0.\tag{9}$$

相应于(7)～(9)的备择命题是：

$$\exists(\rho,\lambda,\mu)\in(W_+^*\times Y_+^*\times Z^*)\backslash\{0\}:$$
$$(\rho\tilde{f}+\lambda g+\mu h)(D)\geqslant 0;\tag{7'}$$

$$\exists(\rho,\lambda)\in(W_+^*\backslash\{0\})\times Y_+^*:(\rho\tilde{f}+\lambda g)(D)\geqslant 0;\tag{8'}$$

$$\exists(\rho,\lambda)\in(W_+^*\backslash\{0\})\times Y_+^*:$$
$$(\rho\tilde{f}+\lambda g)(D\cap H)\geqslant 0.\tag{9'}$$

由得出 4.4.1 的分析同样得到：

6.2.10 定理 设 D 凸,$\bar{x}$ 是问题(1) 的弱有效解. 若命题(7) 与(7′) 两择一,则有 $0 \neq (\bar{\rho},\bar{\lambda},\bar{\mu}) \in W_+^* \times Q \times Z^*$,使

$$L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu}) \in (D-x)^*.$$

应用 3.4.2 或 3.6.8 可使 6.2.10 具体化.

其次,利用 3.6.6 得出(9)与(9′)两择一,可得到 4.4.2 的以下推广.

6.2.11 定理 设 W_+^* 与 Y_+^* 皆有弱*紧凸基,$h(x) = Bx - b$, $R(B)$ 闭;(f,g) 在 H 上为 $W_+ \times Y_+$ - 次类凸,或为 $*$ 拟凸且 $*$ lsc;$\exists \hat{x} \in H$:$\langle Y_+^* \setminus \{0\}, g(\hat{x})\rangle < 0$. 若 $\bar{x}$ 是 VOP(1)$(D = X)$ 的弱有效解,则 $\bar{x}$ 是 K-T 点.

取 $Z = \{0\}$ 得到:

6.2.12 推论 设 W_+^* 与 Y_+^* 皆有弱*紧凸基;(f,g) 为 $W_+ \times Y_+$ - 次类凸;$\exists \hat{x} \in X$:$\langle Y_+^* \setminus \{0\}, g(\hat{x})\rangle < 0$. 若 $\bar{x}$ 是 VOP(3)$(D = X)$ 的弱有效解,则 $\bar{x}$ 是 K-T 点.

若 $\bar{x}$ 是(1) 的局部弱有效解,$N(B) \cap A^{-1}K_g \subset K_M(\bar{x})$,则 $T^{-1}W_-^\circ \cap N(B) \cap A^{-1}K_g = \varnothing$,这意味着问题

$$\exists x \in N(B): Tx \ll 0,\ Ax \in K_g$$

无解. 于是可用 3.6.7 得出(参照 4.4.4):

6.2.13 定理 设 $\overset{\circ}{W}_+ \neq \varnothing$,$K_g$ 与 $R(B)$ 为闭集,Q 有弱*紧凸基;$\exists \hat{x} \in N(B)$:$\langle Q \setminus \{0\}, A\hat{x}\rangle < 0$. 若 $\bar{x}$ 是 VOP(1)$(D = X)$ 的局部弱有效解,$N(B) \cap A^{-1}K_g \subset K_M(\bar{x})$,则 $\bar{x}$ 是 K-T 点.

取 $Z = \{0\}$ 得:

6.2.14 推论 设 $\overset{\circ}{W}_+ \neq \varnothing$,$K_g$ 闭且 $A^{-1}K_g \subset K_G(\bar{x})$;$Q$ 有弱*紧凸基;$\exists \hat{x} \in X$:$\langle Q \setminus \{0\}, A\hat{x}\rangle < 0$. 若 $\bar{x}$ 是问题(3) 的局部弱有效解,则 $\bar{x}$ 是 K-T 点.

若 $W = \mathbf{R}^n$,$\bar{\rho} = (\bar{\rho}_i) \in \mathbf{R}_+^n$ 是乘子,则当 $\rho_i = 0$ 时,在方程 $L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu}) = 0$ 中目标函数 f_i 完全不起作用. 因此能肯定 $\rho \in \overset{\circ}{\mathbf{R}}{}_+^n$ 的结果有重要价值. 下面给出一个这样的结果.

6.2.15 定理([145]) 设 $\dim X<\infty, W=\mathbf{R}^n, Y=\mathbf{R}^m, N_l=\{x\in G: f_j(x)\leqslant f_j(\bar{x})(\forall j\neq l)\}, N=\bigcap_1^n N_l$. 若 $\bar{x}$ 是问题(3)(对于 $D=X$)的有效解,且

$$K\stackrel{\text{def}}{=}(T^{-1}\mathbf{R}_-^n)\cap(A_I^{-1}\mathbf{R}_-^I)\subset\bigcap\overline{\text{co}}\,K(N_l,\bar{x}), \tag{10}$$

则有 $(\bar{\rho},\bar{\lambda})\in\mathring{\mathbf{R}}_+^n\times\mathbf{R}_+^I$,使 $L_x(\bar{x},\bar{\rho},\bar{\lambda})=0$.

证 由 3.4.13,只需证问题

$$\exists z\in X:\ Tz<0,\ A_Iz\leqslant 0 \tag{11}$$

无解. 若 z 是(11)的解,则 $z\in K$. 不妨设 $T_1z<0, T_1=f'_1(\bar{x})$. 由条件(10),$z\in\overline{\text{co}}\,K(N_1,\bar{x})$,取 $z_k=\sum_s\alpha_{ks}z_{ks}\to z(k\to\infty)$,其中 $\alpha_{ks}\geqslant 0,\sum_s\alpha_{ks}=1, z_{ks}\in K(N_1,\bar{x})$;取 $z_{ksv}\to z_{ks}, t_{ksv}\downarrow 0(v\to\infty)$,使 $x_{ksv}\stackrel{\text{def}}{=}\bar{x}+t_{ksv}z_{ksv}\in N_1$,则 $f_l(x_{ksv})\leqslant f_l(\bar{x})(1<l\leqslant n)$. 因 $\bar{x}$ 是问题(3)的有效解,故 $f_1(x_{ksv})\geqslant f_1(\bar{x})$,这推出 $T_1z_{ks}\geqslant 0$,从而 $T_1z\geqslant 0$,得出矛盾. □

注 可证 $K\supset\bigcap_1^n\overline{\text{co}}\,K(N_l,\bar{x})$ 恒成立,因此条件(10)即 $K=\bigcap_1^n\overline{\text{co}}\,K(N_l,\bar{x})$. 关于使(10)成立的具体条件参考[145].

"John 充分条件"4.5.1 几乎不必改动即可用于 VOP:

6.2.16 定理 设 $(\bar{\rho},\bar{\lambda},\bar{\mu})\in W_+^*\times Q\times Z^*$ 满足 $L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})=0$,则以下每个条件推出 $\bar{x}$ 为问题(1)的严格有效解:

(i) $L(\cdot,\bar{\rho},\bar{\lambda},\bar{\mu})$ 在 $\bar{x}$ 关于 M 为 (F,r)- 严格伪凸,$r\geqslant 0$;

(ii) $\bar{\rho}f$ 与 $\bar{\lambda}g+\bar{\mu}h$ 之一在 $\bar{x}$ 关于 M 为 (F,r)- 严格伪凸,另一个为 (F,r')- 拟凸,$r+r'\geqslant 0$;

(iii) $\bar{\rho}f,\bar{\lambda}g,\bar{\mu}h$ 之一在 $\bar{x}$ 关于 M 为 (F,r)- 严格伪凸,另两个分别为 (F,r')- 拟凸与 (F,r'')- 拟凸,$r+r'+r''\geqslant 0$.

以上条件中的"(F,r)-"等可同时改为"η-"或尽行删去.

以下结果与 4.5.3 相当.

6.2.17 定理 设 $(\bar{\rho},\bar{\lambda},\bar{\mu})\in(W_+^*\setminus\{0\})\times Q\times Z^*$ 满足 $L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})=0$,则以下每个条件推出 $\bar{x}$ 为问题(1)的弱有效解:

(i) $L(\cdot,\bar{\rho},\bar{\lambda},\bar{\mu})$ 在 $\bar{x}$ 关于 M 为(F,r)-伪凸,$r \geqslant 0$;

(ii) $\bar{\rho}f$ 与 $\bar{\lambda}g + \bar{\mu}h$ 在 $\bar{x}$ 关于 M 分别为(F,r)-伪凸与(F,r')-拟凸,$r + r' \geqslant 0$;

(iii) $\bar{\rho}f,\bar{\lambda}g,\bar{\mu}h$ 在 $\bar{x}$ 关于 M 分别为(F,r)-伪凸、(F,r') 拟凸与(F,r'')-拟凸,$r + r' + r'' \geqslant 0$.

若 $\bar{\rho} \in \mathring{W}_{+}^{*}$,则以上每个条件推出 $\bar{x}$ 为式(1) 的真有效解.

证 由 4.5.3,当条件(i)~(iii)之一满足时,$\bar{x}$ 是问题

$$\min \bar{\rho}f(x),\ x \in M \tag{12}$$

的最优解,于是定理结论由 6.1.5 推出. □

类似的结果可参考[121].

下面的结果可看作是 4.5.5 的推广.

6.2.18 定理 设 $\dim X < \infty$,则以下每个条件推出 $\bar{x}$ 是问题(1) 的严格局部有效解:

(i) $(T^{*}W_{+}^{*}) \cap (A^{*}Q + R(B^{*}))^{\circ} \neq \varnothing$;

(ii) K_g 闭且 $\exists \rho \in W_{+}^{*}$:$\langle T^{*}\rho,\ (N(B) \cap A^{-1}K_g)\setminus\{0\}\rangle > 0$.

证 若条件(i) 满足,则有 $\bar{\rho} \in W_{+}^{*}$,使 $-T^{*}\bar{\rho} \in (A^{*}Q + R(B^{*}))^{\circ}$,于是 $\bar{x}$ 是问题(12) 的严格局部最优解(4.5.5),从而 $\bar{x}$ 是 VOP(1) 的严格局部有效解(6.1.5). 当条件(ii) 满足时证明是类似的. □

用类似的方法可将 4.5.8 推广于下:

6.2.19 定理 若存在锥 $K \subset X$ 在 $\bar{x}$ 逼近 M,且

$$\sup_{\rho \in W_{+}^{*}} \inf \langle T^{*}\rho,\ K \cap S(X)\rangle > 0, \tag{13}$$

则 $\bar{x}$ 是问题(1) 的严格局部有效解.

证 只需指明:条件(13) 推出 $\exists \rho \in W_{+}^{*}, \beta > 0, \forall z \in K$:$\langle T^{*}\rho, z\rangle \geqslant \beta|z|$,因此可用 4.5.8. □

参考文献:[45,121,123,124,126,127,130,133,137,141,145,147,198].

§ 6.3 非光滑最优性条件

本节对问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D \tag{1}$$

得出 § 4.6 的结果的某些推广. 以下设 $\bar{x} \in M = D \cap G \cap H$, f, g, h 在 $\bar{x}$ 邻近为 Lip.

6.3.1 定理 设 $W = \mathbf{R}^n, Y = \mathbf{R}^m, K \subset F_D(\bar{x})$ 是一非空开凸锥; h 在 $\bar{x}$ 邻近为 C^1 映射, $R(B)$ 为闭集. 若 $\bar{x}$ 是问题(1) 的局部弱有效解, 则有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+^n \times \mathbf{R}_+^l \times Z^*$, $\bar{\rho} = (\bar{\rho}_k)$, $\bar{\lambda} = (\bar{\lambda}_i)$ (以下概如此), 使得

$$\left\{ \sum \bar{\rho}_k \partial f_k(\bar{x}) + \sum \bar{\lambda}_i \partial g_i(\bar{x}) + B^* \bar{\mu} \right\} \cap K^* \neq \varnothing. \tag{2}$$

若 D 凸且 $D^\circ \neq \varnothing$, 则可取 $K = F_D(\bar{x})$, $K^* = (D - \bar{x})^*$; 若 $\bar{x} \in D^\circ$, 则(2)成为

$$0 \in \sum \bar{\rho}_k \partial f_k(\bar{x}) + \sum \bar{\lambda}_i \partial g_i(\bar{x}) + B^* \bar{\mu}. \tag{3}$$

证 不妨设 $g(\bar{x}) = 0, R(B) = Z, \forall u_k \in \partial f_k(\bar{x}), v_i \in \partial g_i(\bar{x})$: $\{u_k : 1 \leqslant k \leqslant n\}$ 与 $\{v_i : 1 \leqslant i \leqslant m\}$ 皆正线性无关(否则式(3) 已不待证). 令

$$K_0 = \{z : f_k(\bar{x}, z) < 0 \quad (1 \leqslant k \leqslant n)\},$$

$$K_1 = \{z : g_i^\circ(\bar{x}, z) < 0 \quad (1 \leqslant i \leqslant m)\},$$

$$K_2 = K, \ K_3 = N(B) = K_H(\bar{x}).$$

则 K_0, K_1, K_2 是非空开凸锥,

$$K_0^* = \sum \mathbf{R}_- \partial f_k(\bar{x}), K_1^* = \sum \mathbf{R}_- \partial g_i(\bar{x})$$

(4.1.8), $K_3^* = R(B^*)$. 若 $z \in X \backslash F_f(\bar{x})$, 则 $\exists t_j \downarrow 0, z_j \to z$, 对某个 k 有 $f_k(\bar{x} + t_j z_j) \geqslant f_k(\bar{x})$, 于是

$$f_k^\circ(\bar{x}, z) \geqslant \limsup_j \frac{f_k(\bar{x} + t_j z) - f_k(\bar{x})}{t_j}$$

$$\geqslant \limsup_{j} \frac{f_k(\bar{x}+t_j z)-f_k(\bar{x}+t_j z_j)}{t_j}=0,$$

故 $z \overline{\in} K_0$. 可见 $K_0 \subset F_f(\bar{x})$. 其次, $K_1 \subset F_G(\bar{x})$, 因此

$$\bigcap_{i=0}^{3} K_i \subset F_f(\bar{x}) \cap F_G(\bar{x}) \cap F_D(\bar{x}) \cap K_H(\bar{x})$$
$$\subset F_f(\bar{x}) \cap K_M(\bar{x}) = \varnothing.$$

余下的证明如同 4.6.1. □

6.3.2 定理([155]) 设 $W=\mathbf{R}^n, Y=\mathbf{R}^m, Z=\mathbf{R}^p$, D 为闭集. 若 $\bar{x}$ 是问题(1)的局部弱有效解, $k>0$ 充分大, 则有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in \mathbf{R}_+^n \times \mathbf{R}_+^l \times \mathbf{R}^p$, 使得

$$0 \in \partial_x L(x, \bar{\rho}, \bar{\lambda}, \bar{\mu}) + k \partial d_D(\bar{x}). \tag{4}$$

若 $\bar{x} \in D^\circ$, 则式(4)成为

$$0 \in \partial_x L(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}). \tag{5}$$

证 不妨设 $f(\bar{x})=0, g(\bar{x})=0$. 取充分小的 $\varepsilon>0$, 令 $f_\varepsilon = f+\varepsilon e, e=(1,1,\cdots,1)$; $\omega_\varepsilon=(f_\varepsilon, g, h)$, $T=(\mathbf{R}_+^n \times \mathbf{R}_+^m \times \mathbf{R}^p) \cap S^{n+m+p-1}$. $\forall t=(\rho,\lambda,\mu) \in T$, 定义

$$\varphi_{t\varepsilon}(x) = \langle t, \omega_\varepsilon(x) \rangle = L(x,t) + \varepsilon \langle \rho, e \rangle.$$

不妨设 $\bar{x}$ 是(1)的全局弱有效解, $\operatorname{Lip} \omega_\varepsilon < k$, 因此 $\operatorname{Lip} \varphi_{t\varepsilon} < \varepsilon$. 于是 $\varphi_\varepsilon(x) = \max_{t \in T} \varphi_{t\varepsilon}(x)$ 有定义且 $\operatorname{Lip} \varphi_\varepsilon < k$.

任给 $x \in D$, 有 $t=(\rho,\lambda,\mu) \in T$, $\rho=(\rho_i)$, $\lambda=(\lambda_i)$, $\mu=(\mu_j)$, 使得

$$\varphi_\varepsilon(x) - \varphi_{t\varepsilon}(x) = \sum \rho_i f_{\varepsilon i}(x) + \sum \lambda_i g_i(x) + \sum \mu_j h_j(x).$$

因 $\bar{x}$ 为弱有效解, 一个初等的讨论可得出, 上式右端 $n+m+p$ 项皆非负. 令

$$\Phi_\varepsilon = (f_{\varepsilon 1}^+, \cdots, f_{\varepsilon n}^+, g_1^+, \cdots, g_m^+, h),$$

则可指明 $\varphi_\varepsilon(x) = \langle t, \Phi_\varepsilon(x) \rangle = |\Phi_\varepsilon(x)|$, $t=\Phi_\varepsilon(x)/|\Phi_\varepsilon(x)|$ 由 x 唯一决定.

直接看出 $\varphi_\varepsilon(\bar{x}) = \sqrt{n}\,\varepsilon \leqslant \inf_D \varphi_\varepsilon(x) + \sqrt{n}\,\varepsilon$. 取 $\delta = \sqrt{n\varepsilon}$, 由

1.6.4 有 $x_\epsilon \in D \cap \overline{B}(\overline{x}, \sqrt{n\epsilon})$，使得

$$\varphi_\epsilon(x_\epsilon) = \min_{x \in D} [\varphi_\epsilon(x) + \sqrt{\epsilon}|x - x_\epsilon|].$$

余下的证明与 4.6.2 一样. □

类似于 4.6.3 引入以下术语：若存在 $(\overline{\rho},\overline{\lambda},\overline{\mu}) \in (W_+^* \setminus\{0\}) \times Q \times Z^*$，使得

$$0 \in \partial f_{\overline{\rho}}(\overline{x}) + g_{\overline{\lambda}}(\overline{x}) + \partial h_{\overline{\mu}}(\overline{x}) \tag{6}$$

$(f_{\overline{\rho}} = \overline{\rho}f$，其余仿此)，或当 $W = \mathbf{R}^n, Y = \mathbf{R}^m, Z = \mathbf{R}^p$ 时有

$$0 \in \sum \overline{\rho}_i \partial f_i(\overline{x}) + \sum \overline{\lambda}_i \partial g_i(\overline{x}) + \sum \overline{\mu}_j \partial h_j(\overline{x}), \tag{7}$$

则称 $\overline{x}$ 为问题(1) 的广义 K-T 点.

以下结果可看作 4.6.4 的推广.

6.3.3 定理 设 $\overline{x} \in D^\circ$ 是问题(1) 的局部弱有效解，$W = \mathbf{R}^n, Y = \mathbf{R}^m$，则以下每个条件推出 $\overline{x}$ 是(1) 的广义 K-T 点：

(i) h 在 $\overline{x}$ 邻近为 C^1 映射，$R(B) = Z$，且存在 $z \in N(B)$，使 $g_i^\circ(\overline{x},z) < 0\ (\forall i \in I)$；

(ii) $g_i(\forall i \in I)$ 在 $\overline{x}$ 关于 M 为广义伪凸，$h(x) = Bx - b$，$R(B) = Z$，且 $\exists \hat{x} \in D: g_I(\hat{x}) \ll 0, h(\hat{x}) = 0$；

(iii) $Z = \mathbf{R}^p, \forall u_i \in \partial g_i(\overline{x}), v_j \in \partial h_j(\overline{x}): \{u_i, v_j : i \in I, 1 \leqslant j \leqslant p\}$ 线性无关；

(iv) $Z = \mathbf{R}^p, \exists z \in X: g_i^\circ(\overline{x},z) < 0\ (\forall i \in I), h_j^\circ(\overline{x}, \pm z) = 0, \forall v_j \in \partial h_j(\overline{x}): \{v_j : 1 \leqslant j \leqslant p\}$ 线性无关.

证 假定条件(i) ~ (iv) 之一满足，则依 6.3.1 或 6.3.2 有 $0 \neq (\overline{\rho},\overline{\lambda},\overline{\mu}) \in \mathbf{R}_+^n \times \mathbf{R}_+^l \times Z^*$，使式(3) 或(5) 成立. 设 $\overline{\rho} = 0$，则$(\lambda, \mu) \neq 0, 0 \in \sum \overline{\lambda}_i \partial g_i(x) + \partial h_{\overline{\mu}}(\overline{x})$. 余下指明这将引出矛盾. 因条件(i) ~ (iv) 如同 4.6.4 中一样，故证法与 4.6.4 并无不同. □

为推广 4.6.5 ~ 4.6.8，设集 K, Φ 分别依 §4.6(12)，(13)；且令

$$\Psi = \bigcup \left\{ \sum \rho_i \partial f_i(\overline{x}) : \rho \in \mathbf{R}_+^n \text{ 且 } |\rho| = 1 \right\}. \tag{8}$$

显然，$\bar{x}$是(1)的广义K-T点 $\Leftrightarrow 0 \in \Psi + \Phi \Leftrightarrow -\Psi \cap \Phi \neq \varnothing$.
基于此，可以建立：

6.3.4 定理 设$W = \mathbf{R}^n$;$\bar{x}$是问题(1)的局部弱有效解. 若Φ为弱*闭凸集且$K \subset K_M(\bar{x})$,Φ,K依§4.6(12),(13)，则$\bar{x}$是VOP(1)的广义K-T点.

证 若$\bar{x}$非广义K-T点，则$(-\Psi) \cap \Phi = \varnothing$. 易见$\Psi$弱*紧凸，因此由1.1.3有$z \in X$，使$-\langle \Psi, z\rangle > \langle \Phi, z\rangle$. 这推出$\langle \Psi, z\rangle < 0$, $\langle \Phi, z\rangle \leqslant 0$. 前者推出$f_l^\circ(\bar{x}, z) < 0 (1 \leqslant l \leqslant n)$，从而$z \in F_f(\bar{x})$(参看6.3.1之证)；后者推出$z \in K \subset K_M(\bar{x})$(参考4.6.5之证). 这得出$z \in F_f(\bar{x}) \cap K_M(\bar{x})$，矛盾于6.2.1. □

6.3.5 定理 设$W = \mathbf{R}^n$,$\bar{x}$是问题(1)的局部弱有效解，Ψ依(8). 则$\Psi \cap T_M^*(\bar{x}) \neq \varnothing$；当$N_M(\bar{x}) \subset \Phi$($\Phi$依§4.6(13))时$\bar{x}$是问题(1)的广义K-T点.

证 若$\Psi \cap T_M^*(\bar{x}) = \varnothing$，则有$z \in X$，使得$\langle \Psi, z\rangle < \langle T_M^*(\bar{x}), z\rangle$. 这推出$z \in F_f(\bar{x})$, $z \in T_M(\bar{x}) \subset K_M(\bar{x})$(参考4.6.6与6.3.4之证)，得出矛盾. 余下的结论是显然的. □

4.6.9有如下推广.

6.3.6 定理 设$W_+^{*\circ} \neq \varnothing$,$Y_+^*$有弱*紧凸基；$h(x) = Bx - b$；$(f, g)$在$H = h^{-1}(0)$上为$W_+ \times Y_+$-次类凸，或为*拟凸且*lsc；$\exists \hat{x} \in H: \langle Y_+^* \setminus \{0\}, g(\hat{x})\rangle < 0$. 若$\bar{x}$是问题(1)(对于$D = X$)的弱有效解，则$\bar{x}$是(1)的广义K-T点.

证 $\bar{x}$是弱有效解意味着问题

$$\exists x \in H: f(x) - f(\bar{x}) \ll 0, g(x) \leqslant 0$$

无解. 由3.6.6有$(\bar{\rho}, \bar{\lambda}) \in (W_+^* \setminus \{0\}) \times Y_+^*$，使得$(\bar{\rho} f + \bar{\lambda} g)(H) \geqslant \bar{\rho} f(\bar{x})$. 如同4.4.2之证，由此推出$\bar{\lambda} \in Q$,

$$\bar{\rho} f(\bar{x}) = (\bar{\rho} f + \bar{\lambda} g)(\bar{x}) = \min\ (\rho f + \lambda g)(H).$$

如同4.6.9之证，由此推出

$$0 \in \partial f_{\bar{\rho}}(\bar{x}) + \partial g_{\bar{\lambda}}(\bar{x}) + R(B^*),$$

这表明$\bar{x}$是广义K-T点. □

应用择一定理亦可得出局部最优性条件. 例如,设 $W=\mathbf{R}^n, Y=\mathbf{R}^m$, $\bar{x}$ 是问题

$$\min f(x), \quad g(x)\leqslant 0 \tag{9}$$

的局部弱有效解,且以下条件满足:

$$\exists z\in X: g_i^\circ(\bar{x},z)<0 \quad (\forall i\in I), \tag{10}$$

则由 6.3.3 知 $\bar{x}$ 是广义 K-T 点. 下面给出一个基于择一定理的证明. 问题

$$\exists x\in X: f_l^\circ(\bar{x},x)<0, g_i^\circ(\bar{x},x)\leqslant 0 \quad (1\leqslant l\leqslant n, i\in I)$$

必无解(否则有 $x\in F_f(\bar{x})\cap K_G(\bar{x})$,参考 4.1.8). 用 3.4.8(注意到条件(10))得出 $(\bar{\rho},\bar{\lambda})\in \mathbf{R}_+^n\setminus\{0\})\times\mathbf{R}_+^l$,使得

$$\sum\bar{\rho}_l f_l^\circ(\bar{x},z)+\sum\bar{\lambda}_i g_i^\circ(\bar{x},z)\geqslant 0\ (\forall z\in X).$$

这恰好意味着(参考 2.2.1,2.3.3)

$$\begin{aligned}0\in\ &\partial_z\Big(\sum\bar{\rho}_l f_l^\circ(\bar{x},z)+\sum\bar{\lambda}_i g_i^\circ(\bar{x},z)\Big)\Big|_{z=0}\\ &\subset\sum\bar{\rho}_l\partial f_l(\bar{x})+\sum\bar{\lambda}_i\partial g_i(\bar{x}),\end{aligned}$$

可见 $\bar{x}$ 是广义 K-T 点.

§4.6 或 §6.2 中的充分条件都可在本节的框架内推广. 例如,4.6.10 或 6.2.16 有如下推广.

6.3.7 定理 设 $(\bar{\rho},\bar{\lambda},\bar{\mu})\in W_+^*\times Q\times Z^*$ 满足(5),则以下每个条件推出 $\bar{x}$ 是问题(1)的严格有效解:

(i) $L(\cdot,\bar{\rho},\bar{\lambda},\bar{\mu})$ 在 $\bar{x}$ 关于 M 为广义(F,r)-严格伪凸,$r\geqslant 0$;

(ii) $\bar{\rho}f$ 与 $\bar{\lambda}g+\bar{\mu}h$ 之一在 $\bar{x}$ 关于 M 为广义(F,r)-严格伪凸,另一个为广义(F,r')-拟凸,$r+r'\geqslant 0$;

(iii) $\bar{\rho}f$,$\bar{\lambda}g$ 与 $\bar{\mu}h$ 之一在 $\bar{x}$ 关于 M 为广义(F,r_1)-严格伪凸,另两个在 $\bar{x}$ 关于 M 分别为广义(F,r_2)-拟凸与广义(F,r_3)-拟凸,$r_1+r_2+r_3\geqslant 0$.

以上 F 如 1.5.5 且是取定的.

证 若 $\bar{x}$ 非严格有效解,则 $\exists x\in M\setminus\{\bar{x}\}: f(x)\leqslant f(\bar{x})$,于

是 $\bar{\rho}f(x) \leqslant \bar{\rho}f(\bar{x}), \bar{\lambda}g(x) \leqslant \bar{\lambda}g(\bar{x}), \bar{\mu}h(x) = 0 = \bar{\mu}h(\bar{x})$,

$$L(x,\bar{\rho},\bar{\lambda},\bar{\mu}) \leqslant L(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu}).$$

由此可如同4.6.10之证，由条件(i)，(ii)，(iii)中任一个引出矛盾.

□

下面考虑 6.2.18 的一个非光滑推广. 设 Φ 依 §4.6(13),

$$\Psi = \bigcup \{\partial f_\rho(\bar{x}) : \rho \in W_+^*\}.$$

6.3.8 定理 设 $\dim X < \infty, \forall (\rho,\lambda,\mu) \in W_+^* \times Q \times Z^*$: $\rho f, \lambda g, \mu h$(分别记作 f_ρ, g_λ, h_μ)在 $\bar{x}$ 正则. 若 $(-\Psi) \cap \Phi^\circ \neq \varnothing$，则 $\bar{x}$ 为问题(1)的严格局部有效解.

证 若定理结论不真，则有 $x_n \in M\backslash\{\bar{x}\}$，使得 $x_n \to \bar{x}, f(x_n) \leqslant f(\bar{x})$. 令 $t_n = |x_n - \bar{x}|, z_n = (x_n - \bar{x})/t_n$，则 $x_n = \bar{x} + t_n z_n$. 不妨设 $z_n \to z \in S(X)$. $\forall (\rho,\lambda,\mu) \in W_+^* \times Q \times Z^*$，有

$$\begin{aligned} f_\rho^\circ(\bar{x},z) &= \lim_n \frac{\rho f(\bar{x}+t_n z) - \rho f(\bar{x})}{t_n} \\ &\leqslant \lim_n \frac{\rho f(\bar{x}+t_n z) - \rho f(x_n)}{t_n} = 0; \\ g_\lambda^\circ(\bar{x},z) &= \lim_n \frac{\lambda g(\bar{x}+t_n z) - \lambda g(\bar{x})}{t_n} \\ &\leqslant \lim_n \frac{\lambda g(\bar{x}+t_n z) - \lambda g(x_n)}{t_n} = 0; \\ h_\mu^\circ(\bar{x},z) &= \lim_n \frac{\mu h(\bar{x}+t_n z) - \mu h(x_n)}{t_n} = 0. \end{aligned}$$

若 $\mu \in \partial g_\lambda(\bar{x}), v \in \partial h_\mu(\bar{x})$，则

$$\langle u+v,z\rangle \leqslant g_\lambda^\circ(\bar{x},z) + h_\mu^\circ(\bar{x},z) \leqslant 0.$$

可见 $-z \in \Phi^*$. 取 $w \in (-\Psi) \cap \Phi^\circ$，则 $w(z) < 0$. 另一方面，设 $-w \in \partial f_\rho(\bar{x}), \rho \in W_+^*$，则

$$-w(z) \leqslant f_\rho^\circ(\bar{x},z) \leqslant 0,$$

得出矛盾.

□

参考文献:[45,53,99,134,155].

§6.4 标 量 化

6.1.5 表明,在一定条件下,一个 VOP 可转化为一个标量最优化问题. 本节在更一般的框架内讨论此课题,基本思想源于[144].

给定 $f: X \to W, \varphi: X \to \mathbf{R}, M \subset X$. 考虑问题

$$\min f(x),\ x \in M \tag{V}$$

与

$$\min \varphi(x),\ x \in M. \tag{S}$$

分别以 E_f 与 WE_f 记问题(V) 的有效解集与弱有效解集,以 E_φ 记问题(S) 的最优解集.

6.4.1 定义([144]) 若 $\forall x, y \in M$: $f(x) \leqslant f(y) \Rightarrow \varphi(x) \leqslant \varphi(y)$, $f(x) < f(y)[f(x) \ll f(y)] \Rightarrow \varphi(x) < \varphi(y)$,则称问题(S) 为(V) 的**标量化**或标量表示[**弱标量化**或弱标量表示].

直接看出,若问题(S) 是(V) 的标量化,则 $E_\varphi \subset E_f$;若(S) 是(V) 的弱标量化,则 $E_\varphi \subset WE_f$,且当(S) 有唯一解时 $E_\varphi \subset E_f$. 若 $E_\varphi = E_f$ 或 $E_\varphi = WE_f$,则可以说问题(V) 与(S) 完全等价.

构成标量化的基本方法是令 $\varphi = u \circ f$, $u: W \to \mathbf{R}$ 是适当的单调函数(即 $a \leqslant b \Rightarrow u(a) \leqslant u(b)$). 若 u 单调且 $a \ll b[a < b] \Rightarrow u(a) < u(b)$,则说 u 严格单调[强单调]. 显然,强单调 $\Rightarrow$ 严格单调 $\Rightarrow$ 单调;当 $W = \mathbf{R}$ 时,强单调 $\Leftrightarrow$ 严格单调. 若 $u \in W^*$,则 u 单调 $\Leftrightarrow u \in W_+^*$, u 严格单调 $\Leftrightarrow u \in W_+^* \setminus \{0\}$, $u \in \mathring{W}_+^* \Rightarrow u$ 强单调,当 W 自反, $\mathring{W}_+^* \neq \varnothing$ 时, u 强单调 $\Leftrightarrow u \in \mathring{W}_+^*$.

6.4.2 命题 问题(S) 是(V) 的标量化[弱标量化]的充要条件是存在强单调[严格单调] 函数 $u: f(M) \to \mathbf{R}$,使得 $\varphi = u \circ f$.

证 设(S) 是(V) 的标量化[弱标量化],则 $f(x) = f(y) \Rightarrow \varphi(x) = \varphi(y)$, 因此由 $u(f(x)) = \varphi(x)$ 确定地定义一函数

$u: f(M) \to \mathbf{R}$,且 $u \circ f = \varphi$. 由 6.4.1 直接看出 u 强单调[严格单调]. 条件的充分性是明显的. □

6.4.2 表明,求问题(V) 的标量化[弱标量化]相当于求一强单调[严格单调]函数 u. 为研究方便,通常还要求 u 具有某些较好的性质,如连续性等. 最简单的选择是取 $u \in \mathring{W}_+^*$ 或 $u \in W_+^* \setminus \{0\}$(参看 6.1.5),但有时需要采用某些构造更复杂的 u,下面举几个例子.

6.4.3 例 1° 取定 $a \in W, b \in W_+^\circ$,定义

$$u(x) = \min \{t : x \leqslant a + tb\}, x \in W.$$

不难验证 $u: W \to \mathbf{R}$ 合理定义且严格单调,于是 u 给出问题(V) 的一个弱标量化:

$$\min \min \{t : f(x) \leqslant a + tb\}, \ x \in M. \qquad (S_{ab})$$

显然, x 是问题(S_{ab}) 的最优解 $\Leftrightarrow \exists t \in \mathbf{R}: (x,t)$ 是问题

$$\min t, \ f(x) \leqslant a + tb, \ x \in M \qquad (F_{ab})$$

的最优解. 注意问题 (F_{ab}) 虽多了一个约束条件,但目标函数极其简单. 这一形式的标量化见[164].

2° 取定 $a = (a_i) \in \mathring{\mathbf{R}}_+^n$,则 $u(x) = \min\limits_i a_i x_i (x = (x_i) \in \mathbf{R}^n)$ 在 $\mathbf{R}^n$ 上严格单调. 因此,当 $W = \mathbf{R}^n$ 时问题 (V)有弱标量化

$$\min \min_i a_i f_i(x), \ x \in M. \qquad (S_a)$$

这一形式的标量化见于[225].

6.4.4 引理 给定 W 的覆盖$\{U(t) : t \in \mathbf{R}\}$. 存在强单调或严格单调函数 $u: W \to \mathbf{R}$ 使得 $U(t) = u^{-1}(t)$ 的充要条件分别为

$$\forall x \in U(t), y \in U(s): x \leqslant y \Rightarrow t \leqslant s, x < y \Rightarrow t < s \quad (C)$$

与 $$\forall x \in U(t), y \in U(s): x \leqslant y \Rightarrow t \leqslant s, x \ll y \Rightarrow t < s \quad (C')$$

证明是直接的.

6.4.5 定理([144]) 设 $W_+^\circ \neq \varnothing$,则存在严格单调的 $u \in C(W)$,使得 $E_\varphi = WE_f, \varphi = u \circ f$.

证 不妨设 $WE_f \neq \varnothing$（否则取 $u \in W_+^* \setminus\{0\}$ 即可，参考 6.1.5）. 令 $A = \overline{f(M) + W_+}$，则由 6.1.1 有

$$f(WE_f) \subset \text{wmin}\, f(M) = \text{wmin}\,[f(M) + W_+] \subset \partial A,$$

故 $\partial A \neq \varnothing$. 取定 $e \in W_+^\circ$，令 $U(t) = \partial A + te$.

1° 证 $\{U(t) : t \in \mathbf{R}\}$ 覆盖 W. 取定 $a \in W$，$\bar{x} \in WE_f$. 若 $t \to \infty$，则 $a - f(\bar{x}) \mp te \in \pm W_+^\circ$，于是

$$a \in f(\bar{x}) + W_+^\circ - te \subset A - te;$$

$$a \in f(\bar{x}) - W_+^\circ + te \subset A^c + te;$$

后者即 $a \,\overline{\in}\, A + te$. 因此可定义

$$t_a = \max\{t : a \in A + te\}, \tag{1}$$

且直接看出 $a \in \partial A + t_a e = U(t_a)$.

2° 验证条件(C′). 设 $a \in U(t)$，$b \in U(s)$. 若 $a \leqslant b$，则

$$b \in a + W_+ \subset \partial A + te + W_+ \subset A + te,$$

于是由(1)有 $t \leqslant t_b$. 必定 $t_b = s$，否则由 $b \in U(s) \cap U(t_b)$ 有 a'，$b' \in \partial A$：$a' - b' = (t_b - s)e$，$a' \ll b'$ 或 $b' \ll a'$，这矛盾于6.1.8. 若 $a \ll b$，则必 $t < s$，否则 $t = s$，同样得出 $a', b' \in \partial A$，使 $a' \ll b'$！

由 6.4.4，有严格单调函数 $u : W \to \mathbf{R}$，使 $U(t) = u^{-1}(t)$. 令 $\varphi = u \circ f$.

3° 证 u 连续. 若 u 不连续，则有 $y \in U(t)$，$y_n \in U(t_n)$，$y_n \to y$，但 $|u(y_n) - u(y)| = |t_n - t| \geqslant \varepsilon > 0$，$\varepsilon$ 与 n 无关. 不妨设 $t_n - t \geqslant \varepsilon$. 设 $y = a + te$，$y_n = a_n + t_n e$，$a, a_n \in \partial A$，则 $n \to \infty$ 时，

$$a - a_n = y - y_n + (t_n - t)e \in W_+^\circ,$$

这与 6.1.8 矛盾.

4° 证 $E_\varphi = WE_f$. 已知 $E_\varphi \subset WE_f$. 若 $\bar{x} \in WE_f$，则 $f(\bar{x}) \in \partial A = U(0) = u^{-1}(0)$，于是 $\varphi(\bar{x}) = 0$. 其次，$\forall x \in M$，有 $f(x) \in A + 0e$，故 $t_{f(x)} \geqslant 0$，从而 $\varphi(x) = t_{f(x)} \geqslant 0$，于是 $\bar{x} \in E_\varphi$. □

若以 E_f 代 WE_f，则解决与 6.4.5 相似的问题要困难些. 下面

是一个有限维结果.

6.4.6 定理([144]) 设 $W = \mathbf{R}^n$, $f(M)$ 与 $f(E_f)$ 为紧集, 则存在强单调的 $u \in C(\mathbf{R}^n)$, 使 $E_\varphi = E_f$, $\varphi = u \circ f$.

证 可设 $E_f \neq \varnothing$(否则取 $u(x) = \langle p, x\rangle$, $p \in \mathring{\mathbf{R}}^n_+$, 参考 6.1.5). 取定 $p \in \mathring{\mathbf{R}}^n_+$, 使 $|p| = 1$. 令 $A = f(E_f)$,

$$K(t) = \{x \in \mathbf{R}^n_+ : \langle p, x\rangle = t\}.$$

因 $f(M)$, A, $K(t)$ 为紧集, $f(M) \cap (A - K(t)) = \varnothing (\forall t > 0)$, 故 $\exists \varepsilon_t > 0$:

$$f(M) \cap [A - K(t) + B(0,\varepsilon_t)] = \varnothing \ (t > 0). \tag{2}$$

可设 $t \to \infty$ 时 $\varepsilon_t \geqslant 2$. 令 $\varepsilon_0 = 0$,

$$B = \mathrm{cl} \bigcup_{t \geqslant 0} \{x \in \mathbf{R}^n : d(x, K(t)) \leqslant \varepsilon_t/2\}.$$

易见 $\mathbf{R}^n_+ \setminus \{0\} \subset B^\circ$; 当 $t \to \infty$ 时 $K(t) + \overline{B}_1(0) \subset B$; $\forall a \in A$: $f(M) \cap (a - B) = \{a\}$. 定义

$$h(t) = \min\{t, \sup\{r : K(t) + \overline{B}_r(0) \subset B\}\}, t \geqslant 0,$$

则 $h(0) = 0$, $\inf\limits_{t \geqslant \tau} h(t) > 0 \ (\forall \tau > 0)$. 不难构造一严格单调增的有界连续函数 $s(t)$, 使得 $0 \leqslant s(t) \leqslant h(t) (t \geqslant 0)$.

令 $C = \bigcup\limits_{t \geqslant 0} [K(t) + \overline{B}(0, s(t))]$. C 是闭集: 设 $x_k = y_k + z_k \to x$, $y_k \in K(t_k)$, $z_k \in \overline{B}(0, s(t_k))$, 则 z_k 有界, 不妨设 $z_k \to z$, 从而 $y_k \to y$. 于是 t_k 有界, 从而不妨设 $t_k \to t$, 因此 $s(t_k) \to s(t)$, $z \in \overline{B}(0, s(t))$, $y \in K(t)$, $x = y + z \in C$. 显然 $\mathbf{R}^n_+ \subset C \subset B$. 其次指明

$$C + (\mathbf{R}^n_+ \setminus \{0\}) \subset C^\circ. \tag{3}$$

设 $t \geqslant 0$, $x \in K(t)$, $y \in \overline{B}(0, s(t))$, $z \in \mathbf{R}^n_+ \setminus \{0\}$. 取 $\tau > t$, 使 $x + z \in K(\tau)$, 则

$$\begin{aligned} x + y + z &\in K(\tau) + \overline{B}(0, s(t)) \\ &\subset K(\tau) + B(0, s(\tau)) \\ &\subset [K(\tau) + \overline{B}(0, s(\tau))]^\circ \subset C^\circ. \end{aligned}$$

任给 $(x, a) \in \mathbf{R}^n \times A$, 令

$$\omega(x,a)=\inf\{t: a-x+tp\in C\}.$$

易见 $\omega: \mathbf{R}^n\times A\to\mathbf{R}$ 连续. 若 $x<y$, 则由式(3) 有

$$a-x+\omega(y,a)p=a-y+\omega(y,a)p+(y-x)$$
$$\in C+(\mathbf{R}_+^n\setminus\{0\})\subset C^\circ.$$

于是有 $\varepsilon>0$, 使 $a-x+[\omega(y,a)-\varepsilon]p\in C$, 因而

$$\omega(x,a)\leqslant\omega(y,a)-\varepsilon<\omega(y,a).$$

这表明 $\omega(x,a)$ 对 x 是强单调的.

因 A 紧, 故可定义 $u(x)=\min\limits_{a\in A}\omega(x,a)$. 易验证 u 连续且强单调. 令 $\varphi=u\circ f$, 则 $E_\varphi\subset E_f$. 任给 $x\in M, a\in A$, 若对某个 $t<0$ 有 $a-f(x)+tp\in C$, 则

$$a-f(x)\in C-tp\subset C+(\mathbf{R}_+^n\setminus\{0\})\subset C^\circ,$$

因此有 $\tau>0$: $a-f(x)\in K(\tau)+\overline{B}(0,s(\tau))$, 从而

$$f(x)\in f(M)\cap[A-K(\tau)+\overline{B}(0,s(\tau))],$$

这将与(2)矛盾. 故

$$\omega(f(x),a)=\inf\{t: a-f(x)+tp\in C\}\geqslant 0,$$

从而 $\varphi(x)\geqslant 0$. 若 $x\in E_f$, 则 $f(x)\in A$, 取 $a=f(x)$ 得 $a-f(x)+0\cdot p=0\in C$, 故 $\varphi(x)=0$, 因此 $x\in E_\varphi$. 故得 $E_\varphi=E_f$. □

参考文献: [45, 102, 144, 164, 224, 225].

§6.5 Lagrange 对偶

本节推广 §5.1 和 §5.2 的某些概念与结果. 由于向量序与标量序有重大差别, 对偶理论的推广有一些难以避免的困难.

6.5.1 定义 设 $L: D\times K\to W$, $(\bar{x},\bar{y})\in D\times K$. 若

$$L(\bar{x},\bar{y})\in\inf L(D,\bar{y})\cap\sup L(\bar{x},K)$$

或
$$L(\bar{x},\bar{y})\in\operatorname{winf} L(D,\bar{y})\cap\operatorname{wsup} L(\bar{x},K),$$

则称 $(\bar{x},\bar{y})$ 为 L 在 $D\times K$ 上的**鞍点**或**弱鞍点**.

若 $W=\mathbf{R}$，则此处所定义的鞍点与弱鞍点就是 §5.1 意义下的鞍点. 6.5.1 中采用 inf 与 sup，而不是见于许多文献的 min 与 max（参考[41,143,210]），看来颇不顺眼，但下面将显示出此定义的某些优越性.

为构成问题

$$\min f(x), g(x) \leqslant 0, h(x)=0, x \in D \tag{1}$$

的 Lagrange 对偶，需定义一个**向量 Lagrange 函数**. 为此，可采用几种不同的方法. 下面的定义主要依据[143]，但有所改进.

取定 $e \in W_+^\circ$. 任给 $\lambda \in Y^*$，$\mu \in Z^*$，令 $\tilde{\lambda}=\lambda(\cdot)e$，$\tilde{\mu}=\mu(\cdot)e$. 记 $L_+(Y,W)=\{\sigma \in L(Y,W): \sigma Y_+ \subset W_+\}$. 取定一凸锥 $S \subset L_+(Y,W)$ 与子空间 $T \subset L(Z,W)$，使之满足条件：

(C) $\forall(\lambda,\mu) \in Y_+^* \times Z^*: \tilde{\lambda} \in S, \tilde{\mu} \in T$.

定义问题(1)的向量 Lagrange 函数为

$$L(x,\sigma,\tau)=f(x)+\sigma g(x)+\tau h(x), \tag{2}$$
$$(x,\sigma,\tau) \in D \times S \times T.$$

$L(x,\sigma,\tau)$ 的定义域依赖于 S,T 的选择. 取 $S=L_+(Y,W)$ 与 $T=L(Z,W)$ 是最大的选择；最小的选择是 $S=\{\tilde{\lambda}: \lambda \in Y_+^*\}$，$T=\{\tilde{\mu}: \mu \in Z^*\}$. 若 $W=\mathbf{R}^n$，$e=(1,1,\cdots,1)$，则对任给 $\lambda=(\lambda_i) \in \mathbf{R}_+^n$，有

$$\tilde{\lambda} g(x)=\left(\sum \lambda_i g_i(x), \cdots, \sum \lambda_i g_i(x)\right) \stackrel{\text{def}}{=} \langle\!\langle \lambda, g(x) \rangle\!\rangle,$$

右端的记号依[210].

若 $x \in M$，则 $L(x,S,T) \leqslant f(x)=L(x,0,0)$. 若 $x \in D \backslash M$，可设 $g(x) \not\leqslant 0$，于是 $\exists \lambda \in Y_+^*: \alpha \stackrel{\text{def}}{=} \lambda g(x)>0$，因此对任给 $b \in W$，当 $t \to \infty$ 时$\stackrel{\text{def}}{=}$有 $L(x,t\tilde{\lambda},0)=f(x)+\alpha t e \gg b$. 这就得出

$$\max L(x,S,T)=\sup L(x,S,T)=\begin{cases} f(x), & x \in M, \\ \varnothing, & x \in D \backslash M. \end{cases}$$

可见，问题(1)相当于

$$\min \sup L(x,S,T), x \in D.$$

于是形成(1)的 Lagrange 对偶为

$$\max \inf L(D,\sigma,\tau), (\sigma,\tau) \in S \times T; \tag{1*}$$

而将相应的“Lagrange 弱对偶”定义为

$$\max \operatorname{winf} L(D,\sigma,\tau), (\sigma,\tau) \in S \times T. \tag{1$_w^*$}$$

6.5.2 定理 以下条件互相等价：

(i) $(\bar{x},\bar{\sigma},\bar{\tau})$ 是 $L(x,\sigma,\tau)$ 在 $D \times (S \times T)$ 上的鞍点；

(ii) $\bar{x} \in M, f(\bar{x}) \in \inf L(D,\bar{\sigma},\bar{\tau}), \bar{\sigma}g(\bar{x}) = 0$；

(iii) $\bar{x}$ 与 $(\bar{\sigma},\bar{\tau})$ 分别为问题(1)与(1*)的有效解且

$$f(\bar{x}) \in \inf L(D,\bar{\sigma},\bar{\tau}).$$

$(\bar{\sigma},\bar{\tau})$ 是(1*)的有效解意味着

$$\inf L(D,\bar{\sigma},\bar{\tau}) \cap \max \left[\bigcup_{(\sigma,\tau)\in S\times T} \inf L(D,\sigma,s)\right] \neq \varnothing.$$

证 (i) ⇒ (ii). 设(i)满足，则必 $\bar{x} \in M$(否则 $\sup L(\bar{x},S,T) = \varnothing$，与 $L(\bar{x},\bar{\sigma},\bar{\tau}) \in \sup L(\bar{x},S,T)$ 矛盾)，于是 $\sup L(\bar{x},S,T) = f(\bar{x})$，从而 $f(\bar{x}) = L(\bar{x},\bar{\sigma},\bar{\tau})$，这推出 $f(\bar{x}) \in \inf L(D,\bar{\sigma},\bar{\tau})$，$\bar{\sigma}g(\bar{x}) = 0$.

(ii) ⇒ (iii). 设(ii)满足. 若 $\bar{x}$ 不是(1)的有效解，则 $\exists x \in M$: $f(x) < f(\bar{x})$. 但 $L(x,\sigma,\bar{\tau}) \leqslant f(x)$，这与 $f(\bar{x}) \in \inf L(D,\bar{\sigma},\bar{\tau})$ 矛盾. 若 $(\bar{\sigma},\bar{\tau})$ 不是问题(1*)的有效解，则有 $(\sigma,\tau) \in S \times T, w \in \inf L(D,\sigma,\tau)$: $w > f(\bar{x})$. 但 $L(\bar{x},\sigma,\tau) \leqslant f(\bar{x})$，这与 $w \in \inf L(D,\sigma,\tau)$ 矛盾.

(iii) ⇒ (i). 设条件(iii)满足，则 $\bar{x} \in M$，于是

$$L(\bar{x},\bar{\sigma},\bar{\tau}) = f(\bar{x}) + \bar{\sigma}g(\bar{x}) \leqslant f(\bar{x}) \in \inf L(D,\bar{\sigma},\bar{\tau}),$$

这推出 $\bar{\sigma}g(\bar{x}) = 0$，因此

$$L(\bar{x},\bar{\sigma},\bar{\tau}) \in \inf (D,\bar{\sigma},\bar{\tau}) \cap \sup L(\bar{x},S,T),$$

即 $(\bar{x},\bar{\sigma},\bar{\tau})$ 是 L 的鞍点. □

6.5.2 推广了 5.2.1(i). 6.5.2 的某种特殊形式见于[41,143,210].

类似地可以证明：

6.5.3 定理 若 $\bar{\sigma}g(\bar{x})=0$，则以下条件互相等价：

(i) $(\bar{x},\bar{\sigma},\bar{\tau})$ 是 L 在 $D\times(S\times T)$ 上的弱鞍点；

(ii) $\bar{x}\in M, f(\bar{x})\in \text{winf } L(D,\bar{\sigma},\bar{\tau})$；

(iii) $\bar{x}$ 与 $(\bar{\sigma},\bar{\tau})$ 分别为问题(1)与(1$_w^*$)的弱有效解且 $f(\bar{x})\in$ winf $L(D,\bar{\sigma},\bar{\tau})$.

如同在 §5.2 中一样，可用择一定理导出对偶定理. 令 $\tilde{f}(x)=f(x)-f(\bar{x})$，考虑备择的一对命题：

$$\exists x\in D:\tilde{f}(x)\ll 0, g(x)\leqslant 0, h(x)=0; \tag{3}$$

$$\exists(\rho,\lambda,\mu)\in(W_+^*\setminus\{0\})\times Y_+^*\times Z^*:$$

$$(\rho\tilde{f}+\lambda g+\mu h)(D)\geqslant 0. \tag{4}$$

6.5.4 定理 设命题(3)与(4)两择一，则 $\bar{x}$ 是(1)的弱有效解 $\Leftrightarrow \exists(\bar{\sigma},\bar{\tau})\in S\times T$：$(\bar{x},\bar{\sigma},\bar{\tau})$ 是 L 在 $D\times(S\times T)$ 上的弱鞍点且 $\bar{\sigma}g(\bar{x})=0$.

证 若 $(\bar{x},\bar{\sigma},\bar{\tau})$ 是 L 的弱鞍点且 $\bar{\sigma}g(\bar{x})=0$，则由 6.5.3 知 $\bar{x}$ 是(1)的弱有效解. 下面设 $\bar{x}$ 是问题(1)的弱有效解，则命题(3)无解，于是(4)有解 (ρ,λ,μ)，$\alpha\overset{\text{def}}{=}\rho(e)>0$. 令 $\bar{\sigma}=\tilde{\lambda}/\alpha, \bar{\tau}=\tilde{\mu}/\alpha$，则 $(\bar{\sigma},\bar{\tau})\in S\times T$，由(4)推出 $\bar{\sigma}g(\bar{x})=\lambda(g(\bar{x}))e=0$. 因

$$\langle\rho,L(\bar{x},\bar{\sigma},\bar{\tau})\rangle=\rho f(\bar{x})\leqslant(\rho f+\lambda g+\mu h)(D)$$
$$=\langle\rho,L(D,\bar{\sigma},\bar{\tau})\rangle,$$

故 $\langle\rho,f(\bar{x})\rangle=\min\langle\rho,L(D,\bar{\sigma},\bar{\tau})\rangle$，从而

$$f(\bar{x})\in \text{wmin } L(D,\bar{\sigma},\bar{\tau})\subset \text{winf } L(D,\bar{\sigma},\bar{\tau})$$

(用 6.1.3, 6.1.1). 由 6.5.3，$(\bar{x},\bar{\sigma},\bar{\tau})$ 是 L 的弱鞍点. □

6.5.4 推广了[41, Th. 2]. 为将 6.5.4 中的"择一性"条件具体化，可用适当的择一定理，例如 3.4.2.

下面是 6.5.4 的一个变种.

6.5.5 定理 设 W 自反，$\mathring{W}_+^*\neq\varnothing$，$f|M$ 为 W_+-类凸[$f(M)+W_+$ 凸闭]；$\forall\rho\in\mathring{W}_+^*$，命题

$$\exists x \in D: \rho\tilde{f}(x) < 0, g(x) \leqslant 0, h(x) = 0 \qquad (5)$$

与 $$\exists (\lambda,\mu) \in Y_+^* \times Z^*: (\rho\tilde{f} + \lambda g + \mu h)(D) \geqslant 0 \qquad (6)$$

两择一. 若 $\bar{x}$ 是问题(1) 的真有效解[有效解], 则有$(\bar{\sigma},\bar{\tau}) \in S \times T$, 使$(\bar{x},\sigma,\bar{\tau})$ 是 L 在 $D \times (S \times T)$ 上的鞍点.

证 由 6.1.5, $\exists \rho \in \mathring{W}_+^*: \rho f(\bar{x}) = \min \rho f(M)$. 这推出(5) 无解, 于是(6) 有解$(\lambda,\mu)$. 令 $\bar{\sigma},\bar{\tau}$ 如 6.5.4 之证, 则$(\bar{\sigma},\bar{\tau}) \in S \times T$, $\bar{\sigma}g(\bar{x}) = 0, \rho f(\bar{x}) = \min \langle \rho, L(D,\bar{\sigma},\bar{\tau})\rangle$. 于是由 6.1.3, 6.1.1 有

$$f(\bar{x}) \in \text{pmin}\, L(D,\bar{\sigma},\bar{\tau}) \subset \inf L(D,\bar{\sigma},\bar{\tau}).$$

其次, $f(\bar{x}) \in \sup L(\bar{x},S,T)$, 因此$(\bar{x},\bar{\sigma},\bar{\tau})$ 是 L 的鞍点. □

注 从以上证明易见

$$f(\bar{x}) \in \min L(D,\bar{\sigma},\bar{\tau}) \cap \max L(\bar{x},S,T),$$

因此$(\bar{x},\bar{\sigma},\bar{\tau})$ 在通常的定义下亦是鞍点. 6.5.5 推广了[143, Th. 3.2]与[210, Th. 3.1, 4.1].

对于问题

$$\min f(x),\ g(x) \leqslant 0,\ x \in D, \qquad (7)$$

其向量 Lagrange 函数为 $L(x,\sigma) = f(x) + \sigma g(x)$, 而 Lagrange 对偶与 Lagrange 弱对偶分别为

$$\max \inf L(D,\sigma),\ \sigma \in S \qquad (7^*)$$

与 $$\max \text{winf}\, L(D,\sigma),\ \sigma \in S. \qquad (7_w^*)$$

结合 6.5.3 与 3.6.6 可得:

6.5.6 定理 设 Y_+^* 有弱* 紧凸基; (f,g) 在 D 上为 $W_+ \times Y_+$ - 次类凸; $\exists \hat{x} \in D: \langle Y_+^* \setminus \{0\}, g(\hat{x})\rangle < 0$. 则 $\bar{x}$ 是问题(7) 的弱有效解 $\Leftrightarrow \exists \sigma \in S: (\bar{x},\bar{\sigma})$ 是 L 在 $D \times S$ 上的弱鞍点且 $\bar{\sigma}g(\bar{x}) = 0 \Leftrightarrow$ 存在 $\bar{\sigma} \in S$, 使 $\bar{x}$ 与 $\bar{\sigma}$ 分别为问题(7) 与(7_w^*) 的弱有效解且 $f(\bar{x}) \in \text{winf}\, L(D,\bar{\sigma})$.

6.5.6 推广了[41, Th. 2], [143, Th. 3.3].

类似地, 结合 6.5.2, 6.5.5 与 3.6.6 得:

6.5.7 定理 设 W 自反, $\mathring{W}_+^* \neq \varnothing$, Y_+^* 有弱* 紧凸基; (f,g)

在 D 上为 $W_+ \times Y_+^-$ 类凸；$\exists \hat{x} \in D$：$\langle Y_+^* \setminus \{0\}, g(\hat{x}) \rangle < 0$. 若 $\bar{x}$ 是问题(7)的真有效解(或有效解且 $f(M) + W_+$ 闭)，则有 $\bar{\sigma} \in S$，使 $(\bar{x}, \bar{\sigma})$ 是 L 在 $D \times S$ 上的鞍点，且 $\bar{\sigma}$ 是问题(7*)的有效解.

6.5.7 推广了[270, Th.4.1], [143, Th.3.2].

6.5.8 例([41]) 考虑以下 VOP：

$$\min x, x \leqslant 0, x \in D = [-1,1] \times 0 \subset \mathbf{R}^2. \tag{8}$$

取 $S = L_+(\mathbf{R}^2, \mathbf{R}^2) = \left\{ \sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \geqslant 0 \right\}$，令 $L(x, \sigma) = x + \sigma x, (x, \sigma) \in D \times S$. 任给 $\sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in S$，有

$$\begin{aligned} L(D,\sigma) &= \{x + \sigma x : |x_1| \leqslant 1, x_2 = 0\} \\ &= \{x_1(a+1, c) : |x_1| \leqslant 1\} = \overline{L(D,\sigma)}; \end{aligned}$$

$$\inf L(D,\sigma) = \min L(D,\sigma) = (-a-1, -c);$$

$$\operatorname{winf} L(D,\sigma) = \begin{cases} (-a-1, -c), & c > 0, \\ [-a-1, a+1] \times 0, & c = 0. \end{cases}$$

因此问题(8)的 Lagrange 对偶为

$$\max (-a-1, -c), \sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in S. \tag{8*}$$

直接看出，问题(8*)有唯一最优值$(-1,0)$，有效解为 $\bar{\sigma} = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} \in S$. 另一方面，问题(8)有唯一有效解 $\bar{x} = (-1,0) \in \inf L(D,\bar{\sigma})$. 由 6.5.2，$(\bar{x}, \bar{\sigma})$ 是 $L(x,\sigma)$ 在 $D \times S$ 的鞍点. 其次，问题(8)的弱有效解集为 $M = [-1, 0] \times 0$；

$$\operatorname{wmax} \bigcup_{\sigma \in S} \operatorname{winf} L(D,\sigma) = \mathbf{R} \times 0,$$

由此看出(8)的 Lagrange 弱对偶有弱有效解 $\bar{\sigma} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in S$. 取 $\bar{x} = (x_1, 0) \in M$，则 $\overline{\sigma x} = 0 \Leftrightarrow a x_1 = 0$. 若此条件满足，则由6.5.3知 $(\bar{x}, \bar{\sigma})$ 为 $L(x,\sigma)$ 的弱鞍点.

参考文献：[41, 42, 45, 53, 137, 143, 199, 210, 211, 216, 224,

235].

§6.6 Rockafellar 对偶

要将§5.4的内容推广于VOP，难点在于建立向量函数的共轭函数与次微分理论. 有几种互有差异的导入法. 下面的导入法在某些方面并非完全令人满意，但它对空间的序结构没有特殊要求，因而显得比较简洁.

以下设 $f: D \to W, F: D \to 2^W$. 取定 $e \in W_+^\circ$ 与子空间 $\Sigma \subset L(X,W)$，使得 $\forall u \in X^*: \tilde{u} \overset{\text{def}}{=\!=} u(\cdot)e \in \Sigma$.

6.6.1 定义 F 的共轭函数 $F^*: \Sigma \to 2^W$ 定义为

$$F^*(\sigma) = \max \bigcup_{x \in D} [\sigma(x) - Fx], \sigma \in \Sigma. \tag{1}$$

任给 $(x,w) \in \text{Gr } F$，称

$$\partial F(x,w) \overset{\text{def}}{=\!=} \{\sigma \in \Sigma : \sigma(x) \in w + F^*(\sigma)\} \tag{2}$$

为 F 在 x 关于 w 的**次微分**，令 $\partial F(x) = \bigcup_{w \in Fx} \partial F(x,w)$. 若 $\forall w \in Fx: \partial F(x,w) \neq \varnothing$，则说 F 在 x 处**次可微**. 令

$$\overline{F}x = \max \bigcup_{\sigma \in \Sigma} [\sigma(x) - F^*(\sigma)]. \tag{3}$$

由式(1)推出，$\forall w \in Fx, v \in F^*(\sigma)$，有 $w + v \not< \sigma(x)$，即对 $x \in D, \sigma \in \Sigma$ 有

$$\sigma(x) \overline{\in} Fx + F^*(\sigma) + (W_+ \setminus \{0\}). \tag{4}$$

式(4)与 Young 不等式(§5.3(2))相当. 由式(2)推出，$0 \in \partial F(x,w) \Leftrightarrow w \in \min F(D)$，$0 \in \partial f(x) \Leftrightarrow f(x) \in \min f(D)$. 后者给出有效点的次微分刻画.

若 $W = \mathbf{R}^n, e = (1,1,\cdots,1) \in \mathbf{R}^n, u \in \mathbf{R}^n$，则由(1)有

$$F^*(\tilde{u}) = \max \bigcup_{x \in D} (\langle\!\langle u,x \rangle\!\rangle - Fx)$$

(参考[210,211]). 对于函数 f，式(1)，(2)分别成为

$$f^*(\sigma) = \max\{\sigma(x) - f(x): x \in D\}, \sigma \in \Sigma$$

与 $$\partial f(x) = \{\sigma \in \Sigma: \sigma(x) \in f(x) + f^*(\sigma)\}.$$

关于 $f^*(\sigma)$ 的其他定义，例如可参考[158].

6.6.2 命题 设 $w \in Fx$，则以下结论成立：

(i) $\partial F(x,w) \neq \varnothing \Leftrightarrow w \in \overline{F}x$.

(ii) F 在 x 处次可微 $\Leftrightarrow Fx \subset \overline{F}x$.

证 (i) 若存在 $\sigma \in \partial F(x,w)$，则 $w \in \sigma(x) - F^*(\sigma)$，由此推出 $w \in \overline{F}x$. 否则有 $s \in \Sigma, v \in F^*(s): w < s(x) - v$，从而 $v < s(x) - w \in s(x) - Fx$，这与 $v \in F^*(s)$ 矛盾. 反之，若 $w \in \overline{F}x$，则有 $\sigma \in \Sigma: w \in \sigma(x) - F^*(\sigma)$，即 $\sigma(x) \in w + F^*(\sigma)$. 于是由式(2)有 $\sigma \in \partial F(x,w)$，因此 $\partial F(x,w) \neq \varnothing$.

(ii) 由(i) 推出. □

将 6.6.2 用到函数 f 得出：$\partial f(x) \neq \varnothing \Leftrightarrow f(x) \in \overline{f}(x)$；当 $W = \mathbf{R}$ 时 $\partial f(x) \neq \varnothing \Leftrightarrow f(x) = \overline{f}(x)$.

现在给出关于 Rockafellar 对偶的一个简单框架. 给定函数 $\varphi(x,y): D \times Y \to W$，考虑一对互相关联的问题：

$$\min \varphi(x,0),\ x \in D; \tag{P}$$

$$\max\left[-\varphi^*(0,\tau)\right],\ \tau \in T, \tag{P*}$$

其中 φ^* 是 φ 的共轭函数，即

$$\varphi^*(\sigma,\tau) = \max\{\sigma(x) + \tau(y) - \varphi(x,y): (x,y) \in D \times Y\}, \tag{5}$$

$(\sigma,\tau) \in \Sigma \times T$，$T$ 是 $L(Y,W)$ 的子空间，满足 $\tilde{\lambda} = \lambda(\cdot)e \in T(\forall \lambda \in Y^*)$. 令

$$S(y) = \min \varphi(D,y),\ y \in Y. \tag{6}$$

则 $S: Y \to 2^W$. 当 $S(y)$ 在 $y = 0$ 处次可微时说问题(P) 稳定.

6.6.3 命题 $0 \in \varphi(\bar{x},0) + \varphi^*(0,\bar{\tau}) \Leftrightarrow (0,\bar{\tau}) \in \partial\varphi(\bar{x},0) \Rightarrow \bar{x}$ 与 $\bar{\tau}$ 分别为问题(P) 与(P*) 的有效解.

证 直接看出 $0 \in \varphi(\bar{x},0) + \varphi^*(0,\bar{\tau}) \Leftrightarrow -\varphi(\bar{x},0) \in \varphi^*(0,\bar{\tau}) \Leftrightarrow (0,\bar{\tau}) \in \partial\varphi(\bar{x},0)$.

下面设 $-\varphi(\bar{x},0)\in\varphi^*(0,\bar{\tau})$，这意味着

$$\varphi(\bar{x},0)\in\min\{\varphi(x,y)-\bar{\tau}(y):(x,y)\in D\times Y\}. \tag{7}$$

若 $\bar{x}$ 非问题(P)的有效解，则 $\exists x\in D:\varphi(x,0)<\varphi(\bar{x},0)$，这矛盾于式(7). 若 $\bar{\tau}$ 非问题(P^*)的有效解，则有 $\tau\in T, w\in\varphi^*(0,\tau)$，使得 $\varphi(\bar{x},0)<-w$，这矛盾于

$$w\in\varphi^*(0,\tau)=\max\{\tau(y)-\varphi(x,y):(x,y)\in D\times Y\}. \quad\square$$

6.6.4 引理 假定：(i) W 自反且 $\mathring{W}_+^*\neq\varnothing$；(ii) W_+ 是正规序锥([93,2.1.2])；(iii) D 为紧集；(iv) $\varphi(x,y)$ 对 x 为 $*$ wlsc，对 y 为 W_+-凸. 则 $S(y)$ 为 W_+-凸，且

$$S^*(\tau)=\varphi^*(0,\tau), \bar{S}(0)=\max[-\varphi^*(0,T)]. \tag{8}$$

证 $\forall y_1,y_2\in Y, t\in J$，令 $y=ty_1+t'y_2$. 由引理所给条件及 6.1.7 有

$$\begin{aligned}&tS(y_1)+t'S(y_2)\subset t\varphi(D,y_1)+t'\varphi(D,y_2)\\&\subset\varphi(D,y)+W_+\subset S(y)+W_+,\end{aligned}$$

可见 $S(y)$ 为 W_+-凸. 其次，用(1)及 6.1.1 有

$$\begin{aligned}\varphi^*(0,\tau)&=\max\{\tau(y)-\varphi(x,y):(x,y)\in D\times Y\}\\&=\max\bigcup_{y\in Y}[\tau(y)-\varphi(D,y)-W_+]\\&=\max\bigcup_{y\in Y}[\tau(y)-\min\varphi(D,y)-W_+]\\&=\max\bigcup_{y\in Y}[\tau(y)-S(y)-W_+]\\&=\max\bigcup_{y\in Y}[\tau(y)-S(y)]=S^*(\tau).\end{aligned}$$

利用已得结果及式(3)立得 $\bar{S}(0)=\max[-\varphi^*(0,T)]$，于是式(8)得证. $\square$

6.6.5 定理 假定6.6.4的条件满足. 若 $\bar{x}$ 是问题(P)的有效解且(P)稳定，则存在问题(P^*)的有效解 $\bar{\tau}$，使得

$$-\varphi(\bar{x},0)\in\varphi^*(0,\bar{\tau}).$$

证 因(P)稳定，故 $S(y)$ 在 $y=0$ 处次可微. 结合 6.6.2 与式(8)得

$$\varphi(\bar{x},0) \in \min \varphi(D,0) = S(0)$$
$$\subset \overline{S}(0) = \max[-\varphi^*(0,T)],$$

因此 $-\varphi(\bar{x},0) \in \min \varphi^*(0,T)$. 这表明存在 $\bar{\tau} \in T$,使 $-\varphi(\bar{x},0) \in \varphi^*(0,\bar{\tau})$,且 $\bar{\tau}$ 是问题(P^*)有效解. □

§6.7 Mond-Weir 对偶与 Wolfe 对偶

设 f,g,h 可微,$L(x,\sigma,\tau)$ 与 S,T 如 §6.5. 考虑互相关联的向量最优化问题

$$\min f(x),\ g(x) \leqslant 0,\ h(x) = 0; \tag{1}$$

$$\begin{cases} \min f(x), \\ L_x(x,\sigma,\tau) = 0, \sigma g(x) + \tau h(x) \geqslant 0, \\ (\sigma,\tau) \in S \times T; \end{cases} \tag{2}$$

$$\begin{cases} \max L(x,\sigma,\tau), \\ L_x(x,\sigma,\tau) = 0, (\sigma,\tau) \in S \times T. \end{cases} \tag{3}$$

(2)与(3)分别称为(1)的 **Mond-Weir 对偶**与 **Wolfe 对偶**. 分别以 M,N,P 记问题(1),(2),(3)的可行集. 若 $\forall x \in M,(y,\sigma,\tau) \in N$,有 $f(x) \not< f(y)$,则说问题(1)与(2)是**弱对偶**的. 当 $W = \mathbf{R}$ 时,这意味着

$$\alpha = \inf_{x\in M} f(x) \geqslant \sup_{(y,\sigma,\tau)\in N} f(y) = \beta.$$

直接看出,若问题(1)与(2)弱对偶,$\bar{x} \in M,(\bar{x},\bar{\sigma},\bar{\tau}) \in N$,则有 $\bar{\sigma}g(\bar{x}) = 0$,且 $\bar{x}$ 与 $(\bar{x},\bar{\sigma},\bar{\tau})$ 分别为问题(1)与(2)的有效解. 类似地,可定义问题(1)与(3)的弱对偶关系,且可指明:若问题(1)与(3)弱对偶,$\bar{x} \in M,(\bar{x},\bar{\sigma},\bar{\tau}) \in P$,则 $\bar{x}$ 与 $(\bar{x},\bar{\sigma},\bar{\tau})$ 分别为问题(1)与(3)的有效解. 可见,给出使问题(1)与(2)或(3)弱对偶的条件有重要意义.

6.7.1 定理 以下每个条件推出问题(1)与(2)弱对偶:

(i) $\forall (x,\sigma,\tau) \in N, \exists \rho \in \mathring{W}^*_+ : \rho L(x,\lambda,\tau)$ 在 x 关于 M 为$(F,$

r)- 伪凸，$r \geqslant 0$；

(ii) $\forall (x,\sigma,\tau) \in N, \exists \rho \in \mathring{W}_{+}^{*}: \rho f$ 与 $\rho(\sigma g + \tau h)$ 在 x 关于 M 分别为(F, r_1)- 严格伪凸与(F, r_2)- 拟凸，$r_1 + r_2 \geqslant 0$；

(iii) $\forall (x,\sigma,\tau) \in N, \exists \rho \in \mathring{W}_{+}^{*}: \rho f$ 与 $\rho(\sigma g + \tau h)$ 在 x 关于 M 分别为(F, r_1) 拟凸与(F, r_2) 严格伪凸，$r_1 + r_2 \geqslant 0$.

以上 F 如 1.5.5 且是给定的(下同).

证 设问题(1)与(2)非弱对偶，即有 $y \in M$，$(x,\sigma,\tau) \in N$，使 $f(y) < f(x)$，从而

$$L(y,\sigma,\tau) \leqslant f(y) < f(x) \leqslant L(x,\sigma,\tau). \tag{4}$$

若条件(i)满足，则由

$$F(y, \rho L_x(x,\sigma,\tau)) + r|y - x| = r|y - x| \geqslant 0$$

推出 $\rho L(y,\sigma,\tau) \geqslant \rho L(x,\sigma,\tau)$，这与(4)矛盾. 若条件(ii)满足，则由 $\rho f(y) < \rho f(x)$ 与 $\sigma g(y) + \tau h(y) \leqslant 0 \leqslant \sigma g(x) + \tau h(x)$ 分别推出

$$F(y, \rho f'(x)) + r_1|y - x| < 0$$

与
$$F(y, \rho(\sigma g'(x) + \tau h'(x))) + r_2|y - x| \leqslant 0.$$

于是

$$\begin{aligned}
0 &= F(y, \rho L_x(x,\sigma,\tau)) \\
&\leqslant F(y, \rho f'(x)) + F(y, \rho(\sigma g'(x) + \tau h'(x))) \\
&< -(r_1 + r_2)|y - x|,
\end{aligned}$$

得出矛盾. 类似地可从条件(iii)引出矛盾. □

类似地可证明：

6.7.2 定理 以下每个条件推出问题(1)与(3)弱对偶：

(i) $\forall (x,\sigma,\tau) \in P, \exists \rho \in \mathring{W}_{+}^{*}: \rho L(\cdot,\sigma,\tau)$ 在 x 关于 M 为(F, r)- 伪凸，$r \geqslant 0$；

(ii) $\forall (x,\sigma,\tau) \in P, \exists \rho \in \mathring{W}_{+}^{*}, \rho f, \rho\sigma g$ 与 $\rho\tau h$ 在 x 关于 M 分别为(F, r_1)- 凸、(F, r_2)- 凸与(F, r_3)- 凸，$r_1 + r_2 + r_3 \geqslant 0$.

类似于本节的结果可参看[56,170,171].其次,6.7.1 与 6.7.2 显然分别推广了 5.7.1 与 5.7.2.以下结果则可看作 5.7.3 的推广.

6.7.3 定理(逆对偶性)　设 $\bar{x}$ 与 (x^*,σ^*,τ^*) 分别为问题(1)与(2)的有效解,且存在 $(\bar{\sigma},\bar{\tau})\in S\times T$ 使 $L_x(\bar{x},\bar{\sigma},\bar{\tau})=0,\bar{\sigma}g(\bar{x})=0$.则以下每个条件推出 $f(\bar{x})\not< f(x^*)\not< f(\bar{x})$ 且 $\bar{x}=x^*$:

(i) $\forall(x,\sigma,\tau)\in N,\forall\rho\in\mathring{W}_+^*(\neq\emptyset)$:$\rho L(\cdot,\sigma,\tau)$ 在 x 关于 M 为 (F,r)-伪凸,而 $\rho L(\cdot,\sigma^*,\tau^*)$ 在 x^* 关于 M 为 (F,r_1) 严格伪凸,$r,r_1\geqslant 0$;

(ii) $\forall(x,\sigma,\tau)\in N,\forall\rho\in\mathring{W}_+^*$:$\rho f$ 与 $\rho(\sigma g+\tau h)$ 在 x 关于 M 分别为 (F,r_1)-伪凸与 (F,r_2)-拟凸,而 ρf 在 x^* 关于 M 为 (F,r)-严格伪凸,$r_1+r_2\geqslant 0,r+r_2\geqslant 0$;

(iii) $\forall(x,\sigma,\tau)\in N,\forall\rho\in\mathring{W}_+^*$:$\rho f$ 与 $\rho(\sigma g+\tau h)$ 在 x 关于 M 分别为 (F,r_1)-拟凸与 (F,r_2)-严格伪凸,$r_1+r_2\geqslant 0$.

证　因条件(i),(ii),(iii)分别蕴涵 6.7.1 之条件(i)~(iii),故问题(1)与(2)弱对偶,因此 $f(\bar{x})\not< f(x^*)$.其次,由 $(\bar{x},\bar{\sigma},\bar{\tau})\in N$ 与 (x^*,σ^*,τ^*) 是(2)的有效解推出 $f(x^*)\not< f(\bar{x})$.下面设 $\bar{x}\neq x^*$.若条件(i)满足,$\rho\in\mathring{W}_+^*$,则由

$$F(x,\rho L_x(x^*,\sigma^*,\tau^*))+r|\bar{x}-x^*|\geqslant 0$$

推出

$$\rho f(\bar{x})\geqslant\rho L(\bar{x},\sigma^*,\tau^*)>\rho L(x^*,\sigma^*,\tau^*)\geqslant\rho f(x^*),$$

这与 $f(x^*)\not< f(\bar{x})$ 矛盾.取定 $\rho\in\mathring{W}_+^*$,使 $\rho f(\bar{x})\leqslant\rho f(x^*)$.若条件(ii)满足,则分别由 $\rho f(\bar{x})\leqslant\rho f(x^*)$ 与 $\sigma^*g(\bar{x})+\tau^*h(\bar{x})\leqslant 0\leqslant\sigma^*g(x^*)+\tau^*h(x^*)$ 推出

$$F(\bar{x},\rho f'(x^*))+r|\bar{x}-x^*|<0$$

与　$$F(\bar{x},\rho(\sigma^*g'(x^*)+\tau^*h'(x^*)))+r_2|\bar{x}-x^*|\leqslant 0.$$

如同 6.7.1 之证，由此推出 $0 < -(r+r_1)|\bar{x}-x^*|$，得出矛盾. 当条件(iii) 满足时同样导致矛盾. □

一个类似的结果是：

6.7.4 定理 设 $(\bar{x},\bar{\sigma},\bar{\tau}) \in M \times S \times T$ 满足 $L_x(\bar{x},\bar{\sigma},\bar{\tau})=0$，$\bar{\sigma}g(\bar{x})=0$；$(x^*,\sigma^*,\tau^*)$ 是问题(3) 的有效解；$\mathring{W}_+^* \neq \varnothing$. 则以下每个条件推出 $\bar{x}=x^*$：

(i) $\forall \rho \in \mathring{W}_+^*$：$\rho L(\cdot,\sigma^*,\tau^*)$ 在 x^* 关于 M 为(F,r)- 严格伪凸，$r \geqslant 0$；

(ii) $\forall \rho \in \mathring{W}_+^*$：$\rho f$，$\rho\sigma^* g$ 与 $\rho\tau^* h$ 在 x^* 关于 M 分别为(F,r_1)-凸、(F,r_2)- 凸与(F,r_3)- 凸，$r_1+r_2+r_3 \geqslant 0$，且其中至少一个是(F,r_i)- 严格凸.

证 设 $\bar{x} \neq x^*$. 因$(\bar{x},\bar{\sigma},\bar{\tau}) \in P$，故

$$L(x^*,\sigma^*,\tau^*) \not< L(\bar{x},\bar{\sigma},\bar{\tau}),$$

于是有 $\rho \in \mathring{W}_+^*$ 使

$$\rho L(x^*,\sigma^*,\tau^*) \geqslant \rho L(\bar{x},\bar{\sigma},\bar{\tau}). \tag{5}$$

若条件(i)满足，则由 $F(\bar{x},\rho L_x(x^*,\sigma^*,\tau^*))+r|\bar{x}-x^*| \geqslant 0$ 推出

$$\begin{aligned}\rho L(x^*,\sigma^*,\tau^*) &< \rho L(\bar{x},\sigma^*,\tau^*) \leqslant \rho f(\bar{x}) \\ &= \rho L(\bar{x},\bar{\sigma},\bar{\tau}),\end{aligned}$$

这与式(5)矛盾. 若条件(ii)满足，则

$$F(\bar{x},\rho f'(x^*)) + r_1|\bar{x}-x^*| \leqslant \rho[f(\bar{x})-f(x^*)];$$

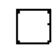
$$F(\bar{x},\rho\sigma^* g'(x^*)) + r_2|\bar{x}-x^*| \leqslant \rho\sigma^*[g(\bar{x})-g(x^*)];$$

$$F(\bar{x},\rho\tau^* h'(x^*)) + r_3|\bar{x}-x^*| \leqslant \rho\tau^*[h(\bar{x})-h(x^*)],$$

且其中至少一个为严格不等式. 于是

$$\begin{aligned}0 &< -(r_1+r_2+r_3)|\bar{x}-x^*| \\ &\quad + \rho L(\bar{x},\sigma^*,\tau^*) - \rho L(x^*,\sigma^*,\tau^*) \\ &\leqslant \rho L(\bar{x},\sigma^*,\tau^*) - \rho L(x^*,\sigma^*,\tau^*)\end{aligned}$$

$$\leqslant \rho L(\bar{x},\bar{\sigma},\bar{\tau}) - \rho L(x^*,\sigma^*,\tau^*),$$

这亦与式(5)矛盾. 

参考文献:[56,99,121,123,142,170,184].

第七章　高阶最优性条件

设 $f: X \to W, g: X \to Y, h: X \to Z, D \subset X$. 对于 VOP

(P)：　$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D$,

本章给出由 f, g, h 的高阶导数表达的最优性条件. 保持上章的记号. 特别, 记 $G = g^{-1}(Y_-), H = h^{-1}(0), \bar{x} \in M = D \cap G \cap H$; $T = f'(\bar{x}), A = g'(\bar{x}), B = h'(\bar{x})$; $L(\cdot, \rho, \lambda, \mu) = \rho f + \lambda g + \mu h$; $K_g = Y_- + \mathbf{R}g(\bar{x}), Q = -K_g^*$; 若 $Y = \mathbf{R}^m$, 则 $I = \{i : g_i(\bar{x}) = 0\}$.

分别以 N_n 与 S_n 记 $\bar{x}$ 满足的某个"n 阶必要条件"与"n 阶充分条件", 这意味着

$$\bar{x} \text{ 是问题(P) 的(局部) 有效解} \Rightarrow \bar{x} \text{ 满足 } N_n;$$

$$\bar{x} \text{ 满足 } S_n \Rightarrow \bar{x} \text{ 是(P) 的(局部) 有效解}.$$

一般说来 $N_n \to N_1, S_1 \to S_n$. 对于仅用一阶条件可以判定的问题, 高阶条件无所为用. 但在相反的情况下, 高阶条件是重要的.

§7.1　二阶条件:光滑情况

本节对 VOP

$$\min f(x),\ g(x) \leqslant 0,\ h(x) = 0 \tag{1}$$

给出二阶最优性条件. 设 $\bar{x} \in M = G \cap H$, f, g, h 在 $\bar{x}$ 邻近为 C^2 函数, $\tilde{f}(x) = f(x) - f(\bar{x})$.

7.1.1 定理　设 $\bar{x}$ 是问题(1) 的 K-T 点(6.2.6), $\bar{\rho}, \bar{\lambda}, \bar{\mu}$ 是 K-T 乘子, $N = \{x \in M : \bar{\rho}\tilde{f}(x) \geqslant \bar{\lambda}g(x) = 0\}$, 则

$$\forall z \in K_N(\bar{x}): L_{xx}(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu})z^2 \geqslant 0. \tag{2}$$

证　设 $z \in K_N(\bar{x})$, 取 $t_k \downarrow 0, z_k \to z$, 使 $x_k \overset{\text{def}}{=} \bar{x} + t_k z_k \in N$, 则

$\bar{\lambda}g(x_k) = \bar{\mu}h(x_k) = 0, \bar{\rho}f(x_k) \geqslant \bar{\rho}f(x)$. 于是

$$0 \leqslant \bar{\rho}\,\bar{f}(x_k) = L(x_k,\bar{\rho},\bar{\lambda},\bar{\mu}) - L(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})$$
$$= \frac{1}{2}t_k^2 L_{xx}(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})z_k^2 + o(t_k^2),$$

故得 $L_{xx}(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})z^2 \geqslant 0$. □

所宜注意者,7.1.1 中并未假定 $\bar{x}$ 是式(1) 的局部有效解. 若 $W = \mathbf{R}$,$\bar{x}$ 是式(1) 的局部最优解,则在 7.1.1 中可取 $N = \{x \in M : \bar{\lambda}g(x) = 0\}$(参考[150,Th. 3. 3]).

7. 1. 2 定理(充分条件) 设 $\dim X < \infty$,$\bar{x}$ 是问题(1) 的 K-T 点,$\bar{\rho},\bar{\lambda},\bar{\mu}$ 是 K-T 乘子;$K = \{z \in N(B) : \langle Q, Az\rangle \leqslant 0 = \langle \bar{\lambda}, Az\rangle\}$.

(i) 若

$$\forall z \in K \cap S(X) : L_{xx}(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})z^2 > 0, \tag{3}$$

则 $\bar{x}$ 是问题(1) 的严格局部有效解.

(ii) 若 $\bar{\rho} \in \mathring{W}_+^*$,且有 $\delta > 0$, 使得

$$\forall (x,z) \in B_\delta(\bar{x}) \times X : d(z, K \cap S(X)) < \delta \Rightarrow L_{xx}(x,\bar{\rho},\bar{\lambda},\bar{\mu})z^2 \geqslant 0, \tag{4}$$

则 $\bar{x}$ 是问题(1) 的局部有效解.

证 (i) 若 $\bar{x}$ 不是(1) 的严格局部有效解,则 $\exists x_k \in M\backslash\{\bar{x}\}$: $x_k \to \bar{x}, f(x_k) \leqslant f(\bar{x})$. 令 $t_k = |x_k - \bar{x}|, z_k = t_k^{-1}(x_k - \bar{x})$,则 $x_k = \bar{x} + t_k z_k, |z_k| = 1$,可设 $z_k \to z \in S(X)$. 由

$$0 \geqslant f(x_k) - f(\bar{x}) = t_k T z_k + o(t_k);$$
$$0 \geqslant \langle Q, g(x_k) - g(\bar{x})\rangle = \langle Q, t_k A z_k + o(t_k)\rangle$$

与

$$0 = h(x_k) - h(\bar{x}) = t_k B z_k + o(t_k),$$

分别得出 $Tz \leqslant 0, \langle Q, Az\rangle \leqslant 0$ 与 $Bz = 0$. 这结合

$$0 = \langle L_x(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu}), z\rangle = \langle \bar{\rho}, Tz\rangle + \langle \bar{\lambda}, Az\rangle$$

得 $\langle \bar{\rho}, Tz\rangle = \langle \bar{\lambda}, Az\rangle = 0$,因此 $z \in K \cap S(X)$. 由

$$0 \geqslant \langle \bar{\rho}, f(x_k) - f(\bar{x})\rangle$$

$$\geqslant L(x_k,\bar{\rho},\bar{\lambda},\bar{\mu}) - L(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})$$

$$= \frac{1}{2}t_k^2 L_{xx}(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})z_k^2 + o(t_k^2)$$

推出 $L_{xx}(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu})z^2 \leqslant 0$,这与条件(3) 矛盾.

(ii) 若 $\bar{x}$ 不是式(1) 的局部有效解,则 $\exists x_k \in M\backslash\{\bar{x}\}: x_k \to \bar{x}$, $f(x_k) < f(\bar{x})$. 如同(i) 之证,可设 $x_k = \bar{x} + t_k z_k, t_k \downarrow 0, z_k \to z \in K \cap S(X)$. 其次,

$$\begin{aligned} 0 &> \langle \bar{\rho}, f(x_k) - f(\bar{x}) \rangle \\ &\geqslant L(x_k,\bar{\rho},\bar{\lambda},\bar{\mu}) - L(\bar{x},\bar{\rho},\bar{\lambda},\bar{\mu}) \\ &= \frac{1}{2}t_k^2 L_{xx}(y_k,\bar{\rho},\bar{\lambda},\bar{\mu})z_k^2, \quad y_k \in [x_k,\bar{x}]. \end{aligned}$$

取 k 充分大,使 $y_k \in B_\delta(\bar{x}), d(z_k, K \cap S(X)) \leqslant |z_k - z| < \delta$,则由条件(4) 有 $L_{xx}(y_k,\bar{\rho},\bar{\lambda},\bar{\mu})z_k^2 \geqslant 0$,得出矛盾. □

注 1 设 N 与 K 分别依 7.1.1 与 7.1.2,则易验证

$$K_N(\bar{x}) \subset K \subset -[N(B) \cap A^{-1}(Y \cap Q^*)].$$

若 $K = \{0\}$,则条件(2),(3) 自动满足. 若 $Y = \mathbf{R}^m$, 则

$$N = \{x \in M : \bar{\rho}\bar{f}(x) \geqslant 0 = \bar{\lambda}_i g_i(x)(\forall i \in I)\};$$

$$K = \{z \in N(B) : A_I z \leqslant 0, \langle \bar{\lambda} g_i'(\bar{x}), z \rangle = 0(\forall i \in I)\}.$$

注 2 若 $W = \mathbf{R}, \bar{\rho} = 1$,则在 7.1.2(i) 之条件下可证 $\exists \varepsilon, \delta > 0, \forall x \in M \cap B_\delta(\bar{x}): f(x) \geqslant f(\bar{x}) + \varepsilon|x - \bar{x}|^2$(参考 4.5.5 与 [150,Th. 5.2]).

注 3 对于例 4.5.6,有 $\bar{\lambda} = a, K = \{0\}$,但 $\bar{x} = 0$ 不是 §4.5(7) 的局部最优解. 可见 7.1.2 中条件 $\dim X < \infty$ 是重要的.

7.1.3 例 考虑有限维问题

$$\begin{cases} \min x_1, \\ (x_1 - 4)^2 + x_2^2 \leqslant 16, (x_1 - 3)^2 + (x_2 - 2)^2 = 13. \end{cases} \tag{5}$$

作 Lagrange 函数

$$\begin{aligned} L = x_1 &+ \lambda[(x_1 - 4)^2 + x_2^2 - 16] \\ &+ \mu[(x_1 - 3)^2 + (x_2 - 2)^2 - 13], \end{aligned}$$

解

$$
\begin{cases}
\partial L/\partial x_1 = 1 + 2\lambda(x_1 - 4) + 2\mu(x_1 - 3) = 0, \\
\partial L/\partial x_2 = 2\lambda x_2 + 2\mu(x_2 - 2) = 0, \\
\lambda[(x_1 - 4)^2 + x_2^2 - 16] = 0, \\
(x_1 - 3)^2 + (x_2 - 2)^2 - 13 = 0, \\
(x_1 - 4)^2 + x_2^2 - 16 \leqslant 0, \lambda \geqslant 0
\end{cases}
$$

得问题(5)的三个 K-T 点：

$$\bar{x} = (0,0), \bar{y} = (32/5, 16/5), \bar{z} = (3 + \sqrt{13}, 2),$$

对应的 K-T 乘子$(\bar{\lambda}, \bar{\mu})$分别为

$$(1/8, 0), (3/40, -1/5) \text{ 与 } (0, -1/2\sqrt{13}).$$

对于点$\bar{x}, \bar{y}$可求得$K = \{0\}$(依 7.1.2 的记号)，因此$\bar{x}, \bar{y}$是(5)的严格局部最优解. 对于点$\bar{z}$有$K = \{(2t, -3t) : t \in \mathbf{R}\}$, $L_{xx}(\bar{z}, \bar{\lambda}, \bar{\mu}) = -1/\sqrt{13}$，条件(3)不满足. 实际上，$\bar{z}$是极大点.

下面是一个类似于 4.5.8 的无限维结果

7.1.4 定理 假定：(i) $\bar{x}$是VOP(1)的K-T点，$\bar{\rho}, \bar{\lambda}, \bar{\mu}$为K-T乘子；(ii) 锥$K \subset X$在$\bar{x}$逼近$M$；(iii) 存$\alpha, \beta > 0$，使得

$$
\begin{cases}
\forall z \in K : \langle T^* \bar{\rho}, z \rangle \leqslant \alpha |z| \\
\Rightarrow L_{xx}(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu}) z^2 \geqslant \beta |z|^2.
\end{cases}
\tag{6}
$$

则$\bar{x}$是VOP(1)的严格局部有效解，且当$W = \mathbf{R}$时$\exists \epsilon, \delta > 0$, $\forall x \in M \cap B_\delta(\bar{x}) : f(x) \geqslant f(\bar{x}) + \epsilon |x - \bar{x}|^2$.

证 不妨设$W = \mathbf{R}, \bar{\rho} = 1$(否则首先考虑标量问题

$$\min \bar{\rho} f(x), x \in M,$$

然后利用 6.1.5(iv) 过渡到VOP(1)). 其次可设$\bar{x} = 0$. 由条件(ii)有$\varphi : M \to K$，使$\varphi(x) = x + o(|x|)(x \to 0)$. 设$x \in M$, $|x|$充分小. 若$\langle f'(0), \varphi(x) \rangle \leqslant \alpha |\varphi(x)|$，则由(6)有

$$
\begin{aligned}
f(x) - f(0) &\geqslant L(x, \bar{\lambda}, \bar{\mu}) \quad L(0, \bar{\lambda}, \bar{\mu}) \\
&= \frac{1}{2} L_{xx}(0, \bar{\lambda}, \bar{\mu}) x^2 + o(|x|^2) \\
&= \frac{1}{2} L_{xx}(0, \bar{\lambda}, \bar{\mu}) [\varphi(x) + o(|x|)]^2 + o(|x|^2)
\end{aligned}
$$

$$\geqslant \frac{\beta}{2}|\varphi(x)|^2 + o(|x|^2)$$
$$= \frac{\beta}{2}|x|^2 + o(|x|^2) \geqslant \frac{\beta}{4}|x|^2.$$

若 $\langle f'(0), \varphi(x)\rangle > \alpha|\varphi(x)|$，则

$$\begin{aligned} f(x) - f(0) &= \langle f'(0), x\rangle + o(|x|) \\ &= \langle f'(0), \varphi(x)\rangle + o(|x|) \\ &> \alpha|\varphi(x)| + o(|x|) \\ &= \alpha|x| + o(|x|) \geqslant \frac{\alpha}{2}|x|^2. \end{aligned}$$

取 $\varepsilon = \min\{\alpha/2, \beta/4\}$，$\delta > 0$ 充分小，则当 $x \in M \cap B_\delta(0)$ 时有 $f(x) \geqslant f(0) + \varepsilon|x|^2$. □

7.1.4 是[150, Th. 5.6]的一个推广.

若 $\dim X < \infty$，取 $K = N(B) \cap A^{-1}K_g$，则 7.1.4 的条件(ii)蕴涵于条件"K_g 闭"(4.5.9)，而条件(iii)可换成如下条件：

$$\forall z \in (K\backslash\{0\}) \cap N(A^*\bar{\lambda}): L_{xx}(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu})z^2 > 0. \tag{7}$$

事实上，若 7.1.4 的条件(iii)不满足，则 $\forall n \geqslant 1, \exists x_n \in K$：

$$\langle T^*\bar{\rho}, x_n\rangle = -\langle\bar{\lambda}, Ax_n\rangle \leqslant n^{-1}|x_n|,$$
$$L_{xx}(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu})x_n^2 < n^{-1}|x_n|^2.$$

令 $z_n = x_n/|x_n|$，可设 $z_n \to z$，则 $0 \neq z \in K$，$L_{xx}(\bar{x}, \bar{\rho}, \bar{\lambda}, \bar{\mu})z^2 \leqslant 0$，$\langle\bar{\lambda}, Az\rangle \geqslant 0$，从而 $\langle\bar{\lambda}, Az\rangle = 0$，$z \in N(A^*\bar{\lambda})$. 这表明条件(7)不满足. 以上分析说明，7.1.4 可看作 7.1.2 的无限维推广.

现在将以上结果用于等式约束问题

$$\min f(x),\ h(x) = 0. \tag{8}$$

设 $\bar{x} \in H$，$L(\cdot, \rho, \mu) = \rho f + \mu h$. 在 7.1.1, 7.1.2 中取 $Y = \{0\}$ 得：

7.1.5 定理 设 $\bar{x}$ 是问题(8)的 K-T 点，$\bar{\rho}, \bar{\mu}$ 是 K-T 乘子.

(i) 若 $N = \{x \in H : \bar{\rho}\tilde{f}(x) \geqslant 0\}$，则

$$L_{xx}(\bar{x}, \bar{\rho}, \bar{\mu})z^2 \geqslant 0\ (\forall z \in K_N(\bar{x})). \tag{9}$$

(ii) 若 $\dim X < \infty$，

$$L_{xx}(\bar{x},\bar{\rho},\bar{\mu})z^2 > 0 \ (\forall z \in N(B) \cap S(X)), \qquad (10)$$

则 $\bar{x}$ 是问题(8) 的严格局部有效解，且当 $W = \mathbf{R}$ 时 $\exists \varepsilon,\delta > 0$, $\forall x \in H \cap B_\delta(\bar{x})$: $f(x) \geqslant f(\bar{x}) + \varepsilon|x - \bar{x}|^2$.

(iii) 若 $\dim X < \infty$, $\bar{\rho} \in \mathring{W}^*_+$, 存在 $\delta > 0$, 使得 $\forall (x,z) \in B_\delta(\bar{x}) \times X$:

$$d(z,N(B) \cap S(X)) < \delta \Rightarrow L_{xx}(x,\bar{\rho},\bar{\mu})z^2 \geqslant 0, \qquad (11)$$

则 $\bar{x}$ 是问题(8) 的局部有效解.

注 若 $W = \mathbf{R}$, $R(B) = Z$, $\bar{x}$ 是问题(8) 的局部最优解，则条件(9) 可写成 $L_{xx}(\bar{x},\bar{\mu})z^2 \geqslant 0 (\forall z \in N(B))$.

7.1.6 推论 设 $W = \mathbf{R}$, $Y = \mathbf{R}^m$, $(A_I,B): X \to \mathbf{R}^I \times Z$ 是满射；$\bar{x}$ 是问题(1) 的局部最优解. 则 $\bar{x}$ 是 K-T 点. 设 $\bar{\lambda},\bar{\mu}$ 是 K-T 乘子，则

$$L_{xx}(\bar{x},\bar{\lambda},\bar{\mu})z^2 \geqslant 0 \ (\forall z \in N(A_I) \cap N(B)).$$

证 由 4.3.3, $\bar{x}$ 是 K-T 点，设 $\bar{\lambda},\bar{\mu}$ 是 K-T 乘子. 令 $\varphi = (g_I, h)$, 则 $\bar{x}$ 亦是问题

$$\min f(x),\ \varphi(x) = 0$$

的局部最优解与 K-T 点，K-T 乘子就是 $(\bar{\lambda}_I,\bar{\mu})$. 因 $R(\varphi'(x)) = \mathbf{R}^I \times Z$, 故由 7.1.5 及其后的注得出：$\forall z \in N(\varphi'(\bar{x})) = N(A_I) \cap N(B)$, 有

$$\begin{aligned} 0 &\leqslant (f + \bar{\lambda}_I g_I + \bar{\mu}h)''(\bar{x})z^2 \\ &= L_{xx}(\bar{x},\bar{\lambda},\bar{\mu})z^2. \end{aligned}$$

□

在 7.1.4 中取 $Y = \{0\}$, $K = N(B)$ 得到：

7.1.7 定理 设 $L(\cdot,\rho,\mu) = \rho f + \mu h$, $(\bar{\rho},\bar{\mu}) \in (W^*_+ \backslash \{0\}) \times Z^*$ 满足 $L_x(\bar{x},\bar{\rho},\bar{\mu}) = 0$; $R(B) = Z$; 存在 $\beta > 0$, 使得

$$\forall z \in N(B): L_{xx}(\bar{x},\bar{\rho},\mu)z^2 \geqslant \beta|z|^2.$$

则 $\bar{x}$ 是 VOP(8) 的严格局部有效解.

证 由 $R(B) = Z$ 推出 $K \stackrel{\text{def}}{=} N(B)$ 在 $\bar{x}$ 逼近 $H = h^{-1}(0)$(1.2.8). 因此 7.1.4 中的条件(i) ~ (iii) 满足. □

参考文献:[9,28,122,134,149,150].

§7.2 二阶条件:非光滑情况

设 $f:\Omega\subset X\to\mathbf{R}$ 为局部Lip. 关于 f 的二阶非光滑最优性条件涉及某种广义二阶方向导数.

7.2.1 定义 1° 若 $x_n\to x,(x_n-x)/|x_n-x|\to a$,则说 $\{x_n\}$ 在方向 a 收敛于 x,记作 $x_n\xrightarrow{a}x$.

2° 任给 $x\in\Omega,a\in S(X)$,令

$$\partial_a f(x)=\{u:\ \text{存在}\ x_n\xrightarrow{a}x,u_n\in\partial f(x_n),\text{使}\ u_n\xrightarrow{*}u\}.\quad(1)$$

3° 任给 $x\in\Omega,a\in S(X),u\in\partial_a f(x)$,称

$$f''_-(x,a,u)=\inf\liminf_n\frac{f(x_n)-f(x)-u(x_n-x)}{t_n^2}\quad(2)$$

为 f 在 x 对 (a,u) 的二阶下方向导数,其中右端外层的 inf 是对所有如下的 t_n,x_n 取的:$t_n\downarrow 0,x_n\to x,t_n^{-1}(x_n-x)\to a,\exists u_n\in\partial f(x_n)$,使 $u_n\xrightarrow{*}u$.

结合(1)与2.2.2(ii)看出 $\partial_a f(x)\subset\partial f(x)$. 若 $f\in C^2$,则易验证 $\partial_a f(x)=f'(x),f''_-(x,a,u)=\frac{1}{2}f''(x)a^2$.

7.2.2 定义 若当 $x_n\xrightarrow{a}x,u_n\in\partial f(x_n)$ 时 $u_n(a)$ 恒收敛,则说 f 在 x 处是半光滑的.

利用2.2.5不难推出:若 f 在 x 半光滑,则7.2.2中的 $u_n(a)$ 收敛于 $\mathrm{D}_+f(x,a)$. C^1 函数显然是半光滑的;可验知连续凸函数是半光滑的.

7.2.3 例 设 X 为 Hilbert 空间,$f(x)=|x|(x\in X)$. 则 $f'(x)=x/|x|(0\neq x\in X),\partial f(0)=\overline{B}_1(0)\subset X$(2.1.2). 任给 $a\in S(X)$. 若 $0\neq x_n\xrightarrow{a}0$,则 $f'(x_n)=x_n/|x_n|\to a$,于是由(1)得

$\partial_a f(0) = \{a\}$. 其次，当 $x_n \xrightarrow{a} 0$ 时恒有

$$(f'(x_n), a) = (x_n/|x_n|, a) \to |a|^2 = 1,$$

可见 f 在 $x = 0$ 是半光滑的. 若 $t_n \downarrow 0, x_n/t_n \to a$，则必 $f'(x_n) = x_n/|x_n| \to a$，于是

$$\liminf_n \frac{f(x_n) - (a, x_n)}{t_n^2} = \liminf_n \frac{|x_n| - (a, x_n)}{t_n^2} \geqslant 0.$$

取 $x_n = t_n a$ 从(2)得出 $f''_-(0, a, a) = 0$.

以下设 $\dim X < \infty$，以 $a \cdot b$ 记 X 中的内积.

7.2.4 定理（必要条件） 假定：(i) 存在 $a \in S(X), \delta > 0, \forall b \in S(X) \cap B_\delta(a), u \in \partial_b f(\bar{x}): u \cdot a \leqslant 0$；(ii) f 在 $\bar{x}$ 是半光滑的；(iii) $\bar{x}$ 是 f 的局部极小点，则 $0 \in \partial_a f(\bar{x}), f''_-(\bar{x}, a, 0) \geqslant 0$.

证 首先证：若闭凸锥 $K \subset X$ 满足 $a \in K^\circ, K \cap S(X) \subset B_\delta(a)$（$a, \delta$ 依条件(i)），则存在 $b \in K \cap S(X)$，使 $0 \in \partial_b f(\bar{x})$. 由条件(iii)，不妨设 $f(B_\delta(\bar{x})) \geqslant f(\bar{x})$. 今指明 $\exists x_n \in B_\delta(\bar{x}) \cap (K + \bar{x})$，使得 $t_n = |x_n - \bar{x}| \to 0$ 且 $f(x_n) = \min f(A_n)$，其中

$$A_n = D_n \cap (K + \bar{x}), D_n = \{x : t_n \leqslant |x - \bar{x}| \leqslant t_{n-1}\}$$

$(n \geqslant 1, t_0 = \delta)$. 分两种情况考虑：

1° 若 $\exists \eta \in (0, \delta)$，$\bar{x}$ 是 f 在 $\bar{B}_\eta(\bar{x}) \cap (K + \bar{x})$ 上的严格极小点，则令 $m = \min\{f(x) : x \in K + \bar{x}$ 且 $|x - \bar{x}| = \eta\}$，

$$S_n = \left\{x \in \bar{B}_\eta(\bar{x}) \cap (K + \bar{x}) : f(x) \leqslant f(\bar{x}) + \frac{m - f(\bar{x})}{2n}\right\}.$$

取 $x_n \in S_n$，使 $t_n \stackrel{\text{def}}{=} |x_n - \bar{x}| = \max\{|x - \bar{x}| : x \in S_n\}$. 易验证 $S_{n+1} \subset S_n \subset B_\eta(\bar{x}), t_n \to 0, f(x_n) = \min f(A_n)(n \geqslant 1)$.

2° 若 1° 的情况不出现，则 $\exists x_n \in K + \bar{x}: f(x_n) = f(\bar{x}), t_n = |x_n - \bar{x}| \to 0$. 显然 x_n 有所要求的性质.

令 $b_n = (x_n - \bar{x})/t_n$. 不妨设 $b_n \to b \in K \cap S(X)$，从而 $x_n \xrightarrow{b} \bar{x}$. 因 $f(x_n) = \min f(A_n)$，由 4.6.6 有 $u_n \in \partial f(x_n) \cap T^*(A_n, x_n)$，于是

$$u_n \in T^*(A_n, x_n) \subset [V(D_n, x_n) \cap T(K+\bar{x}, x_n)]^*$$
$$= V^*(D_n, x_n) + T^*(K+\bar{x}, x_n)$$

(参考 2.5.2,1.4.7). 设 $u_n = v_n + w_n, v_n \in V^*(D_n, x_n), w_n \in T^*(K+\bar{x}, x_n)$. 直接由几何考察看出 $v_n = \tau_n b_n, \tau_n \geqslant 0; w_n \cdot b_n = 0$. 于是

$$u_n \cdot b_n = \tau_n |b_n|^2 + w_n \cdot b_n = \tau_n.$$

不妨设 $u_n \to u \in \partial_b f(\bar{x})$. 由条件(ii)有

$$\mathrm{D}_+ f(\bar{x}, b) = \lim_n u_n \cdot b = \lim_n u_n \cdot b_n = \lim_n \tau_n \geqslant 0.$$

另一方面,由 $\bar{x} + t_n a \in A_n$ 有 $f(\bar{x} + t_n a) \geqslant f(x_n)$, 于是

$$\begin{aligned}\mathrm{D}_+ f(\bar{x}, b) &= \lim_n \frac{f(\bar{x} + t_n b) - f(\bar{x})}{t_n} \\ &= \lim_n \frac{f(\bar{x} + t_n b_n) - f(\bar{x})}{t_n} \\ &= \lim_n \frac{f(x_n) - f(\bar{x})}{t_n} \\ &\leqslant \lim_n \frac{f(\bar{x} + t_n a) - f(\bar{x})}{t_n} \\ &= \mathrm{D}_+ f(\bar{x}, a) \leqslant 0\end{aligned}$$

(用条件(i),(ii)及 2.2.5). 因此 $\tau_n \to \mathrm{D}_+ f(\bar{x}, b) = 0$,这就推出 $u = \lim_n w_n \in T_K^*(0) = K^*$(2.5.2). 因 $a \in K^\circ$,故 $\exists t > 0: a - tu \in K$,于是 $0 \leqslant u \cdot (a - tu) = u \cdot a - t|u|^2 \leqslant -t|u|^2$(用条件(i)), 这推出 $u = 0 \in \partial_b f(\bar{x})$.

选取一列闭凸锥 K_n,使得 $K_{n+1} \subset K_n, a \in K_n^\circ, K_1 \cap S(X) \subset B_\delta(a)$, $\bigcap K_n = \mathbf{R}_+ a$. 由已证结论知,$\forall n \geqslant 1, \exists b_n \in K_n \cap S(X)$: $0 \in \partial_{b_n} f(\bar{x})$. 显然 $b_n \to a$. $\forall n \geqslant 1$, 取 $x_{nk} \xrightarrow{b_n} \bar{x} (k \to \infty), u_{nk} \in \partial f(x_{nk})$,使 $u_{nk} \to 0 (k \to \infty)$;再取 k_n,使 $|x_{nk_n} - \bar{x}| < 1/n, |u_{nk_n}| < 1/n$,且

$$\left| \frac{x_{nk_n} - \bar{x}}{|x_{nk_n} - \bar{x}|} - b_n \right| < \frac{1}{n}.$$

则 $x_{nk_n} \xrightarrow{a} \bar{x}, u_{nk_n} \to 0 \in \partial_a f(\bar{x})$. 直接看出 $f''_-(\bar{x},a,0) \geqslant 0$. □

7.2.5 定理(充分条件) 假定: (i) $\forall a \in S(X), \forall u \in \partial_a f(x): u \cdot a \geqslant 0$; (ii) $\forall a \in S(X): 0 \in \partial_a f(\bar{x}) \Rightarrow f''_-(\bar{x},a,0) > 0$, 则 $\bar{x}$ 是 f 的局部严格极小点.

证 若定理结论不真, 则 $\forall n \geqslant 1, \exists x_n \in \overline{B}_{1/n}(\bar{x}) \setminus \{\bar{x}\}: f(x_n) = \min f(\overline{B}_{1/n}(\bar{x}))$. 令 $a_n = (x_n - \bar{x})/|x_n - \bar{x}|$, 如同7.2.4之证, 由 4.6.6 有 $u_n \in \partial f(x_n), \tau_n \geqslant 0: u_n = -\tau_n a_n$. 可设 $a_n \to a, u_n \to u$. 于是 $x_n \xrightarrow{a} \bar{x}, u \in \partial_a f(\bar{x})$. 由条件(i) 有 $u \cdot a \geqslant 0$. 另一方面,

$$u \cdot a = \lim_n u_n \cdot a_n = -\lim_n \tau_n \leqslant 0,$$

故 $u \cdot a = 0, \tau_n \to 0, u_n \to 0 \in \partial_a f(\bar{x})$. 于是

$$f''_-(\bar{x},a,0) \leqslant \liminf_n \frac{f(x_n) - f(\bar{x})}{|x_n - \bar{x}|^2} \leqslant 0,$$

这与条件(ii)矛盾. □

下面考虑约束最小问题

$$\min f(x),\ g(x) \leqslant 0,\ h(x) = 0, \tag{3}$$

其中 $g: \Omega \to \mathbf{R}^m, h: \Omega \to \mathbf{R}^p$ 与 f 一样为局部 Lip. 以 M 记问题(3)的可行集; 令

$$T = \{t \in \mathbf{R}_+ \times \mathbf{R}_+^m \times \mathbf{R}^p : |t| = 1\};$$

$$L(\cdot,t) = \rho f + \lambda g + \mu h, t = (\rho,\lambda,\mu) \in T;$$

$$T_0 = \{t = (\rho,\lambda,\mu) \in T : 0 \in \partial_x L(\bar{x},t), \lambda g(\bar{x}) = 0\};$$

$$T_a = \{t \in T_0 : 0 \in \partial_a L(\bar{x},t)\};$$

$$D = \{a \in S(X) : \mathrm{D}_+ f(\bar{x},a) \leqslant 0, \mathrm{D}_+ g_i(\bar{x},a) \leqslant 0 (i \in I), \mathrm{D}_+ h_j(\bar{x},a) = 0\}.$$

7.2.6 定理(充分条件) 假定: (i) f, g_i, h_j 在 $\bar{x} \in M$ 是半光滑的; (ii) $\forall t \in T_0: L(\cdot,t)$ 在 $\bar{x}$ 正则; (iii) $\forall a \in D \cap K_M(\bar{x}), T_0 = T_a \Rightarrow \exists t \in T_0: L''_-(\bar{x},t,a,0) > 0$, 则 $\bar{x}$ 是问题(3)的严格局部最优解.

注 $L''_-(\bar{x},t,a,0)$ 记 $L(\cdot,t)$ 在 $\bar{x}$ 对 $(a,0)$ 的二阶下方向导

数.

证 设定理结论不真，于是有 $x_n \in M: x_n \to \bar{x} \neq x_n, f(x_n) \leqslant f(\bar{x})$. 不妨设有 $a \in S(X): x_n \xrightarrow{a} \bar{x}$，容易验证 $a \in D \cap K_M(\bar{x})$.

1° 证 $T_0 = T_a$. 任取 $t = (\rho, \lambda, \mu) \in T_0$；取一列闭凸锥 $K_n \subset X$，使 $a \in K_n^{\circ}, K_{n+1} \subset K_n, \bigcap K_n = \mathbf{R}_+ a$. 取 $\delta_n \downarrow 0$. $\forall m, n \geqslant 1$，当 k 充分大时 $x_k \in \bar{B}(\bar{x}, \delta_m) \cap (K_n + \bar{x}) \overset{\text{def}}{=} S_{mn}$. 易见 $L(x_k, t) \leqslant \rho f(x_k) \leqslant L(\bar{x}, t)$，故有 $x_{mn} \in S_{mn} \setminus \{\bar{x}\}: L(x_{mn}, t) = \min L(S_{mn}, t)$. 固定 n，可设 $b_{mn} \overset{\text{def}}{=} (x_{mn} - \bar{x}) \setminus |x_{mn} - \bar{x}| \to b_n \in K_n \cap S(X) (m \to \infty)$. 如同 7.2.4 之证，有 $u_{mn} \in \partial_x L(x_{mn}, t), u_{mn} = \tau_{mn} b_{mn} + w_{mn}, \tau_{mn} \leqslant 0, w_{mn} \in T^*(K_n + \bar{x}, x_{mn})$. 由此推出 $\tau_{mn} = u_{mn} \cdot b_{mn}$. 可设 $u_{mn} \to u_n \in \partial_x L(\bar{x}, t) (m \to \infty)$. 由条件(i) 有

$$0 \geqslant D_x L(\bar{x}, t, b_n) = \lim_m u_{mn}. b_{mn} = \lim_m \tau_{mn} \overset{\text{def}}{=} \tau_n.$$

由条件(ii)，$D_x L(\bar{x}, t, b_n) = L^{\circ}(\bar{x}, t, b_n) \geqslant 0$，故 $\tau_n = D_x L(\bar{x}, t, b_n) = 0$，从而

$$u_n = \lim_m w_{mn} \in T^*(K_n, 0) = K_n^*.$$

因 $a \in K_n^{\circ}$，故如同 7.2.4 之证有 $u_0 = 0$.

$\forall n \geqslant 1$，取 m_n，使 $|x_{m_n n} - \bar{x}| < 1/n, |b_{m_n n} - b_n| < 1/n, |u_{m_n n}| < 1/n$. 则 $x_{m_n n} \xrightarrow{a} \bar{x}, u_{m_n n} \to 0 \in \partial_a L(\bar{x}, t) (n \to \infty)$，因此 $t \in T_a$.

2° 若 $t \in T_0$ 如上，则由 $L(x_{m_n n}, t) \leqslant L(\bar{x}, t)$ 推出 $L_x''(\bar{x}, t, a, 0) \leqslant 0$，这与条件(iii) 矛盾. □

参考文献：[10,34,36,37].

§7.3 高阶变分集

本节给出切锥概念的一个高阶拓广，即所谓**变分集**，它是定义

变分导数与导出高阶最优性条件的基础.

以下设 $D \subset X, x_i \in X (i \geqslant 0)$ 是给定的. 约定

$$x^n(t,z) = \sum_{j=0}^{n-1} t^j x_j + t^n z; \tag{1}$$

$$V^n(D, x_0, x_1, \cdots, x_{n-1})$$
$$= \{z: \text{当 } t \downarrow 0, x \xrightarrow{D} x_0, w \to z \text{ 时 } x - x_0 + x^n(t,w) \in D\}; \tag{2}$$

$$F^n(D, x_0, x_1, \cdots, x_{n-1})$$
$$= \{z: \text{当 } t \downarrow 0, w \to z \text{ 时 } x^n(t,w) \in D\}; \tag{3}$$

$$T^n(D, x_0, x_1, \cdots, x_{n-1})$$
$$= \{z: \forall t_i \downarrow 0, y_i \xrightarrow{D} x_0, \exists z_i \to z, \text{使 } y_i - x_0 + x^n(t_i, z_i) \in D\}; \tag{4}$$

$$K(D, x_0, x_1, \cdots, x_{n-1})$$
$$= \{z: \exists t_i \downarrow 0, z_i \to z, \text{使 } x^n(t_i, z_i) \in D\}. \tag{5}$$

注意当 $n = 1$ 时,式(2) ~ 式(5) 分别与 § 2.5 中的切锥(1) ~ (4) 重合. 因此,式(2) ~ 式(5) 可看作切锥概念的高阶推广.

下面以 $A^n(D, x_0, x_1, \cdots, x_{n-1})$ 记式(2) ~ 式(5) 中的任一个;当不必指明 x_i 时,就将其简写作 $A^n(D)$, 概称为 **n 阶变分集**. 不难直接验证:

7.3.1 命题 $A^n(\cdot)$ 有以下性质:

(i) $A^n(D, x_0, x_1, \cdots, x_{n-1}) = A^n(D - x_0, 0, x_1, \cdots, x_{n-1})$;

(ii) $\forall \delta > 0: A^n(D) = A^n(D \cap B_\delta(x_0))$;

(iii) $x_0 \in D^\circ \Rightarrow A^n(D) = X, x_0 \overline{\in} \overline{D} \Rightarrow A^n(D) = \varnothing$;

(iv) $\forall \sigma \in \mathrm{GL}(X): \sigma A^n(D, x_0, x_1, \cdots, x_{n-1}) = A^n(\sigma D, \sigma x_0, \sigma x_1, \cdots, \sigma x_{n-1})$;

(v) $\forall \rho > 0: \rho^n A^n(D, x_0, x_1, \cdots, x_{n-1}) = A^n(D, x_0, \rho x_1, \cdots, \rho^{n-1} x_{n-1})$(由此推出当 $n > 1$ 时 $A^n(D)$ 一般不是锥!)

下面的命题汇集了变分集的其他简单性质.

7.3.2 命题 设 $D, E \subset X, x_i \in X$.

(i) $V^n(D)$ 与 $F^n(D)$ 是开集;$T^n(D)$ 与 $K^n(D)$ 是闭集;若 D 凸,则 $A^n(D)(A = V,F,T,K)$ 是凸集.

(ii) $V^n(D) \subset F^n(D) \cap T^n(D)$; $F^n(D) \cup T^n(D) \subset K^n(D)$.

(iii) $F^n(D) = [K^n(D^c)]^c$; $F^n(D \cap E) = F^n(D) \cap F^n(E)$; $K^n(D \cup E) = K^n(D) \cup K^n(E)$;$F^n(D) \cap K^n(E) \subset K^n(D \cap E)$; $V^n(D) \cap T^n(E) \subset T^n(D \cap E)$.

(iv) $A^n(D,x_0,x_1,\cdots,x_{n-1}) \neq \emptyset \Rightarrow x_k \in A^k((D,x_0,x_1,\cdots,x_{n-1}),k = 1,2,\cdots,n-1\}(A = K,T)$.

(v) $x_{n-1} \in A^{n-1}(D,x_0,x_1,\cdots,x_{n-2}) \Rightarrow A^n(D,x_0,x_1,\cdots,x_{n-1}) = X(A = V,F,n \geqslant 2)$.

7.3.3 例 设 $D = \{x \in \mathbf{R}^2 : 0 \leqslant x_1 \leqslant \sqrt{x_2}\}$. 易见

$$V_D(0) = F_D(0) = \mathring{\mathbf{R}}^2_+, T_D(0) = K_D(0) = \mathbf{R}^2_+.$$

其次,较细致的讨论得出,对任给 $x = (x_1,x_2) \in \mathbf{R}$ 有

$$V^2(D,0,x) = \begin{cases} \mathbf{R}^2, & x \in \mathring{\mathbf{R}}^2_+, \\ \mathring{\mathbf{R}}^2_+, & x = 0, \\ \emptyset, & \text{其他情况}; \end{cases}$$

$$F^2(D,0,x) = \begin{cases} \mathbf{R}^2, & x \in \mathring{\mathbf{R}}^2_+, \\ \mathring{\mathbf{R}}^2_+, & x = 0, \\ \mathbf{R} \times (x_1^2,\infty), & x \in (0,\infty) \times \{0\}, \\ (0,\infty) \times \mathbf{R}, & x \in \{0\} \times (0,\infty), \\ \emptyset, & x \in \mathbf{R}^2 \backslash \mathbf{R}^2_+; \end{cases}$$

$$T^2(D,0,x) = \begin{cases} \mathbf{R}^2, & x \in \mathring{\mathbf{R}}^2_+, \\ \mathbf{R}^2_+, & x = 0, \\ \mathbf{R}_+ \times \mathbf{R}, & x \in \{0\} \times (0,\infty), \\ \emptyset, & \text{其他情况}; \end{cases}$$

$$K^2(D,0,x)=\begin{cases}\mathbf{R}^2, & x\in \mathring{\mathbf{R}}^2_+,\\ \mathbf{R}^2_+, & x=0,\\ \mathbf{R}_+\times\mathbf{R}, & x\in\{0\}\times(0,\infty),\\ \mathbf{R}\times[x_1^2,\infty), & x\in(0,\infty)\times\{0\},\\ \varnothing, & x\in\mathbf{R}^2\backslash\mathbf{R}^2_+.\end{cases}$$

以上结果表明，7.3.2(ii)中的包含皆可为真包含；即使 $V_D(0)=F_D(0)$，$T_D(0)=K_D(0)$，但可能$V^2(D,0,x)\neq F^2(D,0,x)$，$T^2(D,0,x)\neq K^2(D,0,x)$. 可见，高阶变分集包含更多的信息.

归纳地定义 $C^n(D,x_0,x_1,\cdots,x_{n-1})$ 如下：

令 $C(D,x_0)=C^1(D,x_0)=\text{cone }(D-x_0)$，

$$C^n(D,x_0,x_1,\cdots,x_{n-1})=C(C^{n-1}(D,x_0,x_1,\cdots,x_{n-2}),x_{n-1}).$$

也将 $C^n(D,x_0,x_1,\cdots,x_{n-1})$ 简写作 $C^n(D)$. 易见，$x_n\in C^n(D)\Leftrightarrow\exists\ t_j>0(1\leqslant j\leqslant n)$：$x_0+\sum t_jx_j\in D$. $C^n(D)$ 总是锥；D 是开(凸)集 $\Rightarrow C^n(D)$ 是开(凸)集.

7.3.4 定理([94]) 设 $x_0\in D\subset X$，$x_1,\cdots,x_{n-1}\in X$. 则

$$F^n(D,x_0,x_1,\cdots,x_{n-1}))\subset C^n(D^\circ,x_0,x_1,\cdots,x_{n-1}). \qquad (6)$$

若 D 凸且 $x_k\in C^k(D,x_0,x_1,\cdots,x_{k-1})(1\leqslant k<n)$，则(6) 为等式. 若 D 凸且 $D^\circ\neq\varnothing$，则$(C^n(D))^\circ=C^n(D^\circ)$.

证 约定 $C^k(D)=C^k(D,x_0,x_1,\cdots,x_{k-1})$，$F^k(D)=F^k(D,x_0,x_1,\cdots,x_{k-1})(1\leqslant k\leqslant n)$.

若 $z\in F^n(D)$，则$\exists\,\delta>0,\forall t\in(0,\delta),w\in B_\delta(z)$：$x^n(t,w)\in D$，这推出 $x^n(t,z)\in D^\circ$，从而 $z\in C^n(D^\circ)$. 可见式(6) 成立.

其次设D凸且$x_k\in C^k(D)(1\leqslant k<n)$，$x_n\in C^n(D^\circ)$. 取 $\tau_{ij}>0(1\leqslant j\leqslant i\leqslant n)$，使得

$$x_0+\sum_{j=1}^{i}\tau_{ij}x_j\in D,\ x_0+\sum_{j=1}^{n}\tau_{nj}x_j\in D^\circ.$$

$\forall t>0$，从方程组

$$\sum_{i=j}^{n} \tau_{ij} s_i = t^j, \quad j = 1,2,\cdots,n$$

可唯一地解出 $s_j = s_j(t)$，且当 $t \downarrow 0$ 时 $0 < s_j(t) = O(t^j)$. 设 $\delta > 0$ 充分小，取定 $t \in (0,\delta)$，$z \in B_\delta(x_n)$，令 $s_j = s_j(t)$，则可设 $s' = 1 - \sum_1^n s_j \in (0,1)$，$x_0 + \sum_{j=1}^{n-1} \tau_{nj} x_j + \tau_{nn} z \in D$. 于是

$$\begin{aligned} x^n(t,z) &= x_0 + \sum_{j=1}^{n-1} \sum_{i=j}^{n} \tau_{ij} s_i x_j + s_n \tau_{nn} z \\ &= s' x_0 + \sum_{i=1}^{n-1} s_i \Big(x_0 + \sum_{j=1}^{i} \tau_{ij} x_j \Big) \\ &\quad + s_n \Big(x_0 + \sum_{j=1}^{n-1} \tau_{nj} x_j + \tau_{nn} z \Big) \in D, \end{aligned}$$

可见 $z \in F^n(D)$. 故式(6)为等式.

最后，设 D 凸，$D^\circ \neq \varnothing$，证 $[C^n(D)]^\circ \subset C^n(D^\circ)$. 只需证 $(C^1(D))^\circ \subset C^1(D^\circ)$. 可设 $x_0 = 0$，令 $E = \text{cone}\, D$，则只要证 $E^\circ \subset \text{cone}\, D^\circ$. 取定 $y = tx \in E^\circ$，$t > 0$，$x \in D$. 设 $\mathbf{R}^+ = (0,\infty)$. 若 $\mathbf{R}^+ x \cap D^\circ = \varnothing$，则由 1.1.2 有 $0 \neq u \in X^*$：$\langle u, \mathbf{R}^+ x\rangle \leqslant \langle u, D\rangle$，这推出 $u(x) \leqslant 0$，$u \in D^*$. 因 $y \in E^\circ$，必 $\exists z \in E$：$u(z) < u(y) \leqslant 0$，这与 $u \in D^* = E^*$ 矛盾. 因此 $\exists s > 0$：$sx \in D^\circ$，从而 $y = ts^{-1} sx \in \text{cone}\, D^\circ$. □

参考文献：[70,92,94,140,148].

§7.4 变分导数

变分导数是表达高阶最优性条件的一个合适工具. n 阶变分导数归纳地定义如下.

7.4.1 定义([92,140]) 设 $x_i \in X (0 \leqslant i \leqslant n)$，$D \subset X$. 约定 $f^{[0]}(D, x_0) = f(x_0)$. 若 $n \geqslant 1$，已定义 i 阶变分导数 $f^{[i]}(D, x_0, x_1,$

$\cdots,x_i),i=0,1,\cdots,n-1,x_n\in K^n(D,x_0,x_1,\cdots,x_{n-1})$，则定义 f 在 x_0 关于$(x_1,x_2,\cdots,x_{n-1})$ 的 n 阶变分导数为

$$f^{[n]}(D,x_0,x_1,\cdots,x_n)=$$

$$\lim_{t\downarrow 0,z\to x_n,x^n(t,z)\in D}\frac{1}{t^n}\Big[f(x^n(t,z))-\sum_{i=0}^{n-1}t^i f^{[i]}(D,x_0,x_1,\cdots,x_i)\Big],\quad (1)$$

只要上式右端极限存在，记号 $x^n(t,z)$ 依 §7.3(1).

式(1) 可等价地表成：当 $t\downarrow 0,z\to x_n,x^n(t,z)\in D$ 时有

$$f(x^n(t,z))=\sum_{i=0}^{n}t^i f^{[i]}(D,x_0,x_1,\cdots,x_i)+o(t^n).\qquad (2)$$

结合(2)与 §7.3(5)得出

$$f^{[n]}\in K^n(f(D),f^{[0]},f^{[1]},\cdots,f^{[n-1]}),\qquad (3)$$

其中 $f^{[i]}=f^{[i]}(D,x_0,\cdots,x_i),0\leqslant i\leqslant n$. 其次，由式(1)归纳地有

$$f^{[n]}(D,x_0,tx_1,\cdots,t^n x_n)=t^n f^{[n]}(D,x_0,x_1,\cdots,x_n).\qquad (4)$$

若 $D\subset E,f^{[n]}(E,x_0,x_1,\cdots,x_n)$ 存在且 $x_n\in K^n(D,x_0,x_1,\cdots,x_{n-1})$，则必 $f^{[n]}(D,x_0,x_1,\cdots,x_n)=f^{[n]}(E,x_0,x_1,\cdots,x_n)$.

若 $D=X$，则在变分导数记号中省去 D，此时 7.4.1 中的条件 $x_n\in K^n(D,x_0,x_1,\cdots,x_{n-1})$ 与式(1) 中的 $x^n(t,z)\in D$ 自动满足，式(1) 简化成

$$f^{[n]}(x_0,x_1,\cdots,x_n)$$

$$=\lim_{t\downarrow 0,z\to x_n}\frac{1}{t_n}\Big[f(x^n(t,z))-\sum_{i=0}^{n-1}t^i f^{[i]}(x_0,x_1,\cdots,x_i)\Big].\qquad (5)$$

特别，取 $n=1$ 得出(参考[8,Def. 2.1])

$$f^{[1]}(x_0,x_1)=\lim_{t\downarrow 0,z\to x_1}\frac{f(x_0+tz)-f(x_0)}{t}.\qquad (6)$$

由(6)看出，若 $f^{[1]}(x_0,x_1)$ 存在，则

$$D_+f(x_0,x_1)=f^{[1]}(x_0,x_1).$$

反之，若 $D_+f(x_0,x_1)$ 存在且 f 在 x_0 邻近为 Lip，则 $f^{[1]}(x_0,x_1)=D_+f(x_0,x_1)$；若 f 在 x_0 可微，则 $f^{[1]}(x_0,\cdot)=f'(x_0)$.

若 $f^{[n]}(x_0,x_1,\cdots,x_n)$ 存在且 $x_n\in K^n(D,x_0,x_1,\cdots,x_{n-1})$，则

必有 $f^{[n]}(D, x_0, x_1, \cdots, x_n) = f^{[n]}(x_0, x_1, \cdots, x_n)$.

下面给出变分导数与 Fréchet 导数的关系.

7.4.2 定理([140]) 若 f 在 x_0 邻近为 C^n 函数,$x(t) = \sum_0^n t^i x_i, n \geqslant 1$,则

$$f^{[n]}(x_0, x_1, \cdots, x_n) = \frac{1}{n!}\frac{\mathrm{d}^n}{\mathrm{d}t^n} f(x(t))\bigg|_{t=0}$$

$$= \sum_{q=1}^{n} \sum_{i_1 + \cdots + i_q = n} \frac{1}{q!} f^{(q)}(x_0) x_{i_1} x_{i_2} \cdots x_{i_q}. \tag{7}$$

证 对 n 用归纳法. 当 $n = 1$ 时式(7)显然成立. 下面设 $n \geqslant 1$ 并作归纳假设. 当 $t \downarrow 0, z \to x_n$ 时,由 Taylor 公式(1.2.4)及归纳假设有

$$\begin{aligned} f(x^n(t,z)) &= f(x(t) + t^n(z - x_n)) \\ &= f(x(t)) + o(t^n) \\ &= \sum_{i=0}^{n} \frac{t^i}{i!}\frac{\mathrm{d}^i}{\mathrm{d}t^i} f(x(t))\bigg|_{t=0} + o(t^n) \\ &= \sum_{i=0}^{n-1} t^i f^{[i]}(x_0, x_1, \cdots, x_i) + \frac{t^n}{n!}\frac{\mathrm{d}^n}{\mathrm{d}t^n} f(x(t))\bigg|_{t=0} + o(t^n), \end{aligned}$$

与式(2)比较得出式(7)的第一个等号. 其次,令 $\varphi(t) = f'(x(t))$,则 $[f(x(t))]' = \varphi(t)x'(t)$. 用高阶 Leibniz 公式及归纳假设,经某些计算后得出:

$$\begin{aligned} &\frac{1}{n!}\frac{\mathrm{d}^n}{\mathrm{d}t^n} f(x(t))\bigg|_{t=0} \\ &= \frac{1}{n!}[\varphi(t)x'(t)]^{(n-1)}\bigg|_{t=0} \\ &= \frac{1}{n!}\sum_{i=0}^{n-1} \binom{n-1}{i} \varphi^{(i)}(0) x^{(n-i)}(0) \\ &= \sum_{i=0}^{n-1} \frac{n-i}{i!\,n} \varphi^{(i)}(0) x_{n-i}, \end{aligned}$$

$$= \sum_{i=1}^{n-1} \frac{n-i}{n} \frac{1}{i!}\varphi^{(i)}(0)x_{n-i} + f'(x_0)x_n$$

$$= \sum_{i=1}^{n-1} \frac{n-i}{n} \sum_{q=1}^{i} \sum_{i_1+\cdots+i_q=i} \frac{1}{q!} f^{(q+1)}(x_0)x_{i_1}\cdots x_{i_q}x_{n-i}$$

$$+ f'(x_0)x_n$$

$$= \sum_{q=1}^{n-1} \sum_{i=q}^{n-1} \frac{n-i}{n} \sum_{i_1+\cdots+i_q=i} \frac{1}{q!} f^{(q+1)}(x_0)x_{i_1}\cdots x_{i_q}x_{n-i}$$

$$+ f'(x_0)x_n$$

$$= \sum_{q=2}^{n} \sum_{i=q}^{n} \frac{n-i+1}{n} \sum_{i_1+\cdots+i_{q-1}=i-1} \frac{1}{(q-1)!} \times$$

$$f^{(q)}(x_0)x_{i_1}\cdots x_{i_{q-1}}x_{n-i+1} + f'(x_0)x_n$$

$$= \sum_{q=1}^{n} \sum_{i_1+\cdots+i_q=n} \frac{1}{q!} f^{(q)}(x_0)x_{i_1}x_{i_2}\cdots x_{i_q}. \qquad \square$$

公式(7)可用来计算变分导数.例如,在式(7)中依次取 $n=1$, 2,3 得到:

$$f^{[1]}(x_0,x_1) = f'(x_0)x_1; \tag{8}$$

$$f^{[2]}(x_0,x_1,x_2) = f'(x_0)x_2 + \frac{1}{2}f''(x_0)x_1^2; \tag{9}$$

$$f^{[3]}(x_0,x_1,x_2,x_3) = f'(x_0)x_3 + f''(x_0)x_1x_2 + \frac{1}{6}f'''(x_0)x_1^3. \tag{10}$$

一般地,由式(7)有

$$f^{[n]}(x_0,x_1,\cdots,x_n) = f'(x_0)x_n + f^{(n)}(x_0,x_1,\cdots,x_{n-1}), \tag{11}$$

其中 $f^{(n)}(x_0,x_1,\cdots,x_{n-1}) = f^{[n]}(x_0,\cdots,x_{n-1},0)$ 与 x_n 无关.

7.4.3 命题 设 f 在 x_0 邻近为 C^1 函数,$f^{[n]}(x_0,\cdots,x_{n-1},0)$ 存在,则公式(11)对任何 $x_n \in X$ 成立.

证 令 $z_t = \sum_0^{n-1} t^i x_i + t^n(z-x_n)$,则当 $t\downarrow 0, z\to x_n$ 时,

$$f(x^n(t,z)) = f(z_t + t^n x_n)$$

$$= f(z_t) + \int_0^1 f'(z_t + st^n x_n) t^n x_n \mathrm{d}s$$

$$= \sum_{i=0}^{n-1} t^i f^{[i]}(x_0, x_1, \cdots, x_i) + t^n f^{[n]}(x_0, \cdots, x_{n-1}, 0)$$

$$+ t^n f'(x_0) x_n + o(t^n),$$

这与式(2)比较得出式(11). □

7.4.3 推广了[148,Prop. 3.1]. 顺便指出,[148] 的证明是不可行的.

7.4.4 命题(链规则) 设 $\varphi = g \circ f$, 则

$$\varphi^{[n]}(D, x_0, x_1, \cdots, x_n) = g^{[n]}(f(D), f^{[0]}, f^{[1]}, \cdots, f^{[n]}), \quad (12)$$

只要上式右端存在,其中 $f^{[i]} = f^{[i]}(D, x_0, x_1, \cdots, x_i)(0 \leqslant i \leqslant n)$.

证 当 $t \downarrow 0, z \to x_n$ 时,由(2)有

$$\varphi(x^n(t,z)) = g\Big(\sum_{i=0}^{n} t^i f^{[i]} + o(t^n)\Big)$$

$$= \sum_{i=0}^{n} t^i g^{[i]}(f(D), f^{[0]}, f^{[1]}, \cdots, f^{[i]}) + o(t^n),$$

由此得出式(12). □

对于具体的函数,计算高阶变分导数往往不易. 看一个简单例子.

7.4.5 例([148]) 设 $F: J \times \mathbf{R}^n \to \mathbf{R}$ 满足条件:(i) $F(\cdot, x)$ 可测;(ii) $F(t, \cdot) \in C^2$;(iii) 存在 $\delta > 0, x_0 \in H^{1,1}(J, \mathbf{R}^n), \beta \in L^1(J)$, 使得

$$|x - x_0(t)| < \delta \Rightarrow |F_x(t,x)| + |F_{xx}(t,x)| \leqslant \beta(t).$$

定义 $f: H^{1,1}(J, \mathbf{R}^n) \to \mathbf{R}$ 为

$$f(x) = \int_0^1 F(t, x(t)) \mathrm{d}t.$$

1° 容易验证 f 可微,且

$$f'(x_0)x_1 = f^{[1]}(x_0, x_1) = \int_0^1 F_x(t, x_0(t)) x_1(t) \mathrm{d}t.$$

2° 设 $\lambda \downarrow 0$, 在 $H^{1,1}(J, \mathbf{R}^n)$ 中 $z \to x_2$. 由(5)有

$$
\begin{aligned}
&f^{[2]}(x_0,x_1,x_2)\\
&=\lim \frac{1}{\lambda^2}[f(x_0+\lambda x_1+\lambda^2 z)-f(x_0)-\lambda f^{[1]}(x_0,x_1)]\\
&=\lim \frac{1}{\lambda^2}\int_0^1[F(t,x_0(t)+\lambda x_1(t)+\lambda^2 z(t))-F(t,x_0(t))\\
&\qquad\qquad -\lambda F_x(t,x_0(t))x_1(t)]\mathrm{d}t\\
&=\lim\int_0^1[F_x(t,x_0(t))z(t)\\
&\qquad\qquad +\frac{1}{2}F_{xx}(t,x_0(t))(x_1(t)+\lambda z(t))^2+\cdots\Big]\mathrm{d}t\\
&=\int_0^1\Big[F_x(t,x_0(t))x_2(t)+\frac{1}{2}F_{xx}(t,x_0(t))x_1^2(t)\Big]\mathrm{d}t\\
&=f'(x_0)x_2+\frac{1}{2}\int_0^1 F_{xx}(t,x_0(t))x_1^2(t)\mathrm{d}t.
\end{aligned}
$$

利用条件(i)～(iii),以上推演可以做得很严格.

参考文献:[8,70,92,140,148].

§7.5 可行集的变分集

本节将§4.1的结果部分地推广到变分集. 设 $D\subset X, x_i\in X(i\geqslant 0)$,记

$$S^n(D,x_0,x_1,\cdots,x_{n-1})=D+\text{span}\{x_0,x_1,\cdots,x_{n-1}\}. \quad (1)$$

7.5.1 引理([140]) 设 $K\subset X$ 为凸锥,$x_0\in K$,

$$x_i\in S^i(K,x_0,x_1,\cdots,x_{i-1}),1\leqslant i\leqslant n.$$

(i) 若 $\alpha_i(\cdot)\in C(\mathbf{R}),\alpha_i(0)>0(0\leqslant i\leqslant n)$,则当 $t\downarrow 0$ 时有 $\beta_n(t)\overset{\text{def}}{=}\sum_0^n t^i\alpha_i(t)x_i\in K$.

(ii) 若 $x_n\in S^n(K^\circ,x_0,x_1,\cdots,x_{n-1})$,则当 $t\downarrow 0$, $z\to x_n$ 时有 $x^n(t,z)\in K^\circ$.

证 (i) 当 $t \downarrow 0$ 时显然 $\beta_0(t) = \alpha_0(t)x_0 \in K$. 设 $x_n \in K + \sum_0^{n-1} r_i x_i$,则当 $t \downarrow 0$ 时,

$$\begin{aligned}\beta_n(t) &\in \sum_{i=0}^{n-1} t^i \alpha_i(t) x_i + t^n \alpha_n(t)\Big(K + \sum_{i=0}^{n-1} r_i x_i\Big) \\ &\subset K + \sum_{i=0}^{n-1} t^i [\alpha_i(t) + t^{n-i} r_i \alpha_n(t)] x_i \\ &= K + \sum_{i=0}^{n-1} t^i \bar{\alpha}_i(t) x_i,\end{aligned}$$

其中 $\bar{\alpha}_i(t) = \alpha_i(t) + t^{n-i} r_i \alpha_n(t) \in C(\mathbf{R})$, $\bar{\alpha}_i(0) = \alpha_i(0) > 0 (0 \leqslant i < n)$. 于是用一归纳假设得出 $\beta_n(t) \in K$.

(ii) 设 $x_n \in K^\circ + \sum_0^{n-1} r_i x_i, r_i \in \mathbf{R}$,则当 $t \downarrow 0, z \to x_n$ 时,利用(i)的结论得出

$$\begin{aligned}x^n(t,z) &= \sum_{i=0}^{n} t^i x_i + o(t^n) \\ &\in \sum_{i=0}^{n-1} t^i x_i + t^n\Big(K^\circ + \sum_{i=0}^{n-1} r_i x_i\Big) + o(t^n) \\ &\subset K^\circ + \sum_{i=0}^{n-1} t^i (1 + t^{n-i} r_i) x_i \\ &\subset K^\circ + K \subset K^\circ.\end{aligned}$$

□

下面约定 $g^{[i]} = g^{[i]}(D, x_0, x_1, \cdots, x_i)$,

$$K_i = Y_- + \sum_{j=0}^{i} \mathbf{R} g^{[j]},\ O_i = Y_-^\circ + \sum_{j=0}^{i} \mathbf{R} g^{[j]}, \tag{2}$$

$K^n(D) = K^n(D, x_0, x_1, \cdots, x_{n-1})$, $F^n(D)$ 仿此. 直接看出 K_i 是凸锥, O_i 是开凸锥; $O_i \subset K_i^\circ$; $K_0 = K_g$(令 $\bar{x} = x_0$);

$$Q_i \stackrel{\text{def}}{=} -K_i^* = -O_i^* = Y_+^* \cap \{g^{[j]} : 0 \leqslant j \leqslant i\}^\perp; \tag{3}$$

$Q_0 = Q$(参照 §4.1).

7.5.2 定理 设 $x_0 \in G \stackrel{\text{def}}{=} g^{-1}(Y_+)$, $g^{[i]} (0 \leqslant i \leqslant n)$ 存在.

(i) 若 $x_n \in K^n(D \cap G)$，K_{n-1} 为闭集，则 $g^{[n]} \in K_{n-1}$.

(ii) 若 $g^{[i]} \in K_{i-1}(1 \leqslant i < n)$，$g^{[n]} \in O_{n-1}$，则

$$x_n \in K^n(D \cap G).$$

(iii) 若 $g^{[i]} = 0(0 \leqslant i < n)$，$x_n \in K^n(D \cap G)$，则 $g^{[n]} \leqslant 0$.

证 (i) 取 $t_k \downarrow 0, z_k \to x_n(k \to \infty)$，使 $x^n(t_k, z_k) \in D \cap G$，则

$$\begin{aligned} 0 &\geqslant \langle Q_{n-1}, g(x^n(t_k, z_k)) \rangle \\ &= \langle Q_{n-1}, \sum_{i=0}^{n} t_k^i g^{[i]} + o(t_k^n) \rangle \\ &= \langle Q_{n-1}, t_k^n g^{[n]} + o(t_k^n) \rangle, \end{aligned}$$

这推出 $g^{[n]} \in -Q_{n-1}^* = K_{n-1}^{**}$. 因 K_{n-1} 是闭凸锥，故有 $g^{[n]} \in K_{n-1}$(1.4.3).

(ii) 由 $g^{[n]}$ 存在知 $x_n \in K^n(D)$，故有 $t_k \downarrow 0, z_k \to x_n(k \to \infty)$，使 $x^n(t_k, z_k) \in D$. 于是结合 §7.4(2) 与 7.5.1(ii) 有

$$g(x^n(t_k, z_k)) = \sum_{i=0}^{n} t_k^i g^{[i]} + o(t_k^n) \in Y_-^\circ,$$

因此 $x^n(t_k, z_k) \in D \cap G$，从而 $x_n \in K^n(D \cap G)$.

(iii) 设 t_k, z_k 如(i)之证，则

$$0 \geqslant g(x^n(t_k, z_k)) = t_k^n g^{[n]} + o(t_k^n),$$

由此得出 $g^{[n]} \leqslant 0$. □

7.5.3 定理 设 $x_0 \in G$，$g^{[i]}(0 \leqslant i \leqslant n)$ 存在.

(i) 若 $g^{[i]} \in K_{i-1}(1 \leqslant i < n)$，$g^{[n]} \in O_{n-1}$，$x_n \in F^n(D)$，则 $x_n \in F^n(D \cap G)$.

(ii) 若 $\varphi \overset{\text{def}}{=} g^{[n]}(D, x_0, x_1, \cdots, x_{n-1}, \cdot)$ 凸且 $\exists \hat{x} \in X: \varphi(\hat{x}) \in O_{n-1}$，$x_n \in F^n(D \cap G)$，则 $g^{[n]} \in O_{n-1}$.

(iii) 若 $g^{[i]} \in K_{i-1}(1 \leqslant i < n)$，$\varphi$ 凸且 $\exists \hat{x} \in X: \varphi(\hat{x}) \in O_{n-1}$，则

$$F^n(D \cap G) = \{x \in F^n(D) : \varphi(x) \in O_{n-1}\}. \tag{4}$$

特别，若 $g^{[i]} = 0(0 \leqslant i < n)$，$\varphi$ 凸且 $\exists \hat{x}: \varphi(\hat{x}) \ll 0$，则

$$F^n(D\cap G)=\{x\in F^n(D):\varphi(x)\ll 0\}. \tag{5}$$

证 (i) 若 $t\downarrow 0, z\to x_n$,则 $x^n(t,z)\in D$,于是如 7.5.2(ii) 之证,可利用 7.5.1(ii) 得出 $g(x^n(t,z))\leqslant 0$,从而 $x_n\in F^n(D\cap G)$.

(ii) 取充分小的 $\tau>0$,使 $x_\tau \stackrel{\text{def}}{=} (1+\tau)x_n-\tau\hat{x}\in F^n(D\cap G)$. 若 $g^{[n]}=\varphi(x_n)\overline{\in} O_{n-1}$,则由 1.1.2 有 $\lambda\in Y^*\setminus\{0\}:\langle\lambda,\varphi(x_n)\rangle\geqslant\langle\lambda,O_{n-1}\rangle$,这推出 $\langle\lambda,\varphi(x_n)\rangle\geqslant 0,\lambda\in -O_{n-1}^*=Q_{n-1}$. 取 $y\in Y_-^\circ$,使 $\varphi(\hat{x})\in Y+\sum_0^{n-1}g^{[i]}$,则当 $0<t\downarrow 0$ 时,有

$$\begin{aligned}
0&\leqslant\langle\lambda,\ (1+\tau)t^n\varphi(x_n)-g(x^n(t,x_\tau))\rangle\\
&=\langle\lambda,\ (1+\tau)t^n\varphi\left(\frac{1}{1+\tau}x_\tau+\frac{\tau}{1+\tau}\hat{x}\right)\\
&\qquad-\sum_{i=0}^{n-1}t^ig^{[i]}-t^n\varphi(x_\tau)+o(t^n)\rangle\\
&\leqslant\langle\lambda,\ \tau t^n\varphi(\hat{x})+o(t^n)\rangle\\
&=\langle\lambda,\ \tau t^n y+o(t^n)\rangle<0,
\end{aligned}$$

得出矛盾. 故必 $g^{[n]}\in O_{n-1}$.

结合(i)与(ii),即得出(iii). □

关于以上两定理可参考[92,140].

7.5.4 推论 设 $x_0\in G, g^{[1]}(D,x_0,x_1)$ 存在.

(i) 若 $x_1\in K_{D\cap G}(x_0)$ 且 $K_g\stackrel{\text{def}}{=}Y_-+\mathbf{R}g(x_0)$ 为闭集,则 $g^{[1]}(D,x_0,x_1)\in K_g$.

(ii) 若 $g^{[1]}(D,x_0,x_1)\in Y_-^\circ+\mathbf{R}g(x_0)$[且 $x_1\in F_D(x_0)$],则 $x_1\in K_{D\cap G}(x_0)$[$x_1\in F_{D\cap G}(x_0)$].

(iii) 若 $g^{[1]}(D,x_0,\cdot)$ 凸且 $\exists\hat{x}\in X:g^{[1]}(D,x_0,\hat{x})\in Y_-^\circ+\mathbf{R}g(x_0)$, $x_1\in F_{D\cap G}(x_0)$,则 $g^{[1]}(D,x_0,x_1)\in Y_-^\circ+\mathbf{R}g(x_0)$.

在 7.5.4 中取 $D=X$ 且设 $A=g'(x_0)$ 存在得出:当 K_g 闭时 $K_G(x_0)\subset A^{-1}K_g$;$K\stackrel{\text{def}}{=}A^{-1}(Y_-^\circ+\mathbf{R}g(x_0))\subset F_G(x_0)$;当 $K\neq\varnothing$ 时 $K=F_G(x_0)$. 这些结论已见于 4.1.2 与 4.1.4. 可见,7.5.4 是

4.1.2 与 4.1.4 的高阶推广.

关于 $H=h^{-1}(0)$ 的变分集的表示要用到较强的可微性条件.

7.5.5 定理([92]) 设 $x_0\in H$,则以下结论成立:

(i) $h^{[i]}(H,x_0,x_1,\cdots,x_i)=0(0\leqslant i\leqslant n)$,只要这些导数存在.

(ii) 若 h 在 x_0 处 n 次可微,在 x_0 邻近可微,$B=h'(x_0)$,$R(B)=Z$,$h^{[i]}(x_0,x_1,\cdots,x_i)=0(1\leqslant i<n)$,则

$$K^n(H,x_0,x_1,\cdots,x_{n-1})=\{x_n:h^{[n]}(x_0,\cdots,x_{n-1},x_n)=0\}. \tag{6}$$

证 (i) 由 $h^{[n]}(H,x_0,x_1,\cdots,x_n)$ 存在知 $x_n\in K^n(H,x_0,x_1,\cdots,x_{n-1})$,于是 $\exists t_k\downarrow 0,z_k\to x_n(k\to\infty)$,使 $y_k\overset{\text{def}}{=\!=}x^n(t_k,z_k)\in H$. 因此

$$0=h(y_k)=\sum_{i=0}^{n}t_k^ih^{[i]}(H,x_0,x_1,\cdots,x_i)+o(t_k^n),$$

由此推出 $h^{[i]}(H,x_0,x_1,\cdots,x_i)=0(0\leqslant i\leqslant n)$.

结论(ii)的证明从略(参考[92,Th.4.3]). □

§7.6 高阶必要条件

本节对 VOP

$$\min f(x),g(x)\leqslant 0,h(x)=0,x\in D \tag{1}$$

给出用变分导数描述的**高阶最优性条件**.

取定 $x_i\in X(0\leqslant i<n)$,$x_0\in M=D\cap G\cap H$. 约定

$$f^{[i]}=f^{[i]}(D,x_0,x_1,\cdots,x_i)(0\leqslant i\leqslant n),$$

$g^{[i]},h^{[i]}$ 仿此;将 §7.5(2) 所定义的 K_i,O_i 记作 $K_i(g),O_i(g)$,令

$$K_i(f)=W_-+\sum_{j=1}^{i}\mathbf{R}f^{[j]},$$

$$O_i(f) = W_-^\circ + \sum_{j=1}^{i} \mathbf{R} f^{[j]}; \qquad (2)$$

$$Q_i(f) = - K_i^*(f), Q_i(g) = - K_i^*(g).$$

假定 $W_+^\circ \neq \varnothing \neq Y_+^\circ$. 令

$$\begin{aligned} F_f^n(D) &= F_f^n(D, x_0, x_1, \cdots, x_{n-1}) \\ &= \{x: \text{当 } 0 < t \downarrow 0, z \to x \text{ 且} \\ &\qquad x^n(t,z) \in D \text{ 时 } f(x^n(t,z)) \ll f(x_0)\}. \end{aligned} \qquad (3)$$

$F_f^n(D)$ 显然是开集, $F_f^1(D) = F_f^n(D, x_0, 0, \cdots, 0)$; 当 $D = X$ 时后者就是 §6.2(2) 所定义的 $F_f(x_0)$. 下面给出的 $F_f^n(D)$ 的变分导数表示与 $F^n(D)$ 的表示(7.5.3)颇相类似.

7.6.1 定理 设 $f^{[i]}(0 \leqslant i \leqslant n)$ 存在.

(i) 若 $f^{[i]} \in K_{i-1}(f)(1 \leqslant i < n)$, $f^{[n]} \in O_{n-1}(f)$, 则 $x_n \in F_f^n(D)$.

(ii) 若 $\varphi \stackrel{\text{def}}{=} f^{[n]}(D, x_0, \cdots, x_{n-1}, \cdot)$ 凸, 且 $\exists \hat{x} \in X: \varphi(\hat{x}) \in O_{n-1}(f)$, $x_n \in F_f^n(D)$, 则 $f^{[n]} \in O_{n-1}(f)$.

(iii) 若 $f^{[i]} \in K_{i-1}(f)(1 \leqslant i < n)$, φ 凸且 $\exists \hat{x} \in X: \varphi(\hat{x}) \in O_{n-1}(f)$, 则

$$F_f^n(D) = \{x_n: f^{[n]}(D, x_0, x_1, \cdots, x_{n-1}, x_n) \in O_{n-1}(f)\}. \quad (4)$$

特别, 若 $f^{[i]} = 0(1 \leqslant i < n)$, φ 凸且 $\exists \hat{x}: \varphi(\hat{x}) \ll 0$, 则

$$F_f^n(D) = \{x: f^{[n]}(D, x_0, x_1, \cdots, x_{n-1}, x) \ll 0\}. \qquad (5)$$

证 (i) 若 $0 < t \downarrow 0, z \to x_n, x^n(t,z) \in D$, 则由所给条件以及 7.5.1(ii) 有

$$f(x^n(t,z)) - f(x_0) = \sum_{i=1}^{n} t^i f^{[i]} + o(t^n) \in W_-^\circ,$$

可见 $x_n \in F_f^n(D)$.

(ii) 取充分小的 $\tau > 0$, 使 $x_\tau \stackrel{\text{def}}{=} (1+\tau)x_n - \tau\hat{x} \in F_f^n(D)$. 若 $f^{[n]} = \varphi(x_n) \bar{\in} O_{n-1}(f)$, 则存在 $\rho \in W^* \setminus \{0\}$, 使得 $\langle \rho, \varphi(x_n) \rangle \geqslant \langle \rho, O_{n-1}(f) \rangle$, 因此 $\langle \rho, \varphi(x_n) \rangle \geqslant 0$.

$$\rho \in -O_{n-1}^*(f) = W_+^* \cap \{f^{[i]} : 1 \leqslant i < n\}^{\perp}.$$

取 $w \in W^\circ$ 使 $\varphi(\hat{x}) - w \in \mathrm{span}\{f^{[i]} : 1 \leqslant i < n\}$. 于是当 $0 < t \downarrow 0$ 时有

$$0 \leqslant \langle \rho, (1+\tau)t^n\varphi(x_n) + f(x_0) - f(x^n(t, x_\tau))\rangle$$

$$\leqslant \langle \rho, t^n\varphi(x_\tau) + \tau t^n\varphi(\hat{x}) - \sum_{i=1}^{n-1} t^i f^{[i]} - t^n\varphi(x_\tau) + o(t^n)\rangle$$

$$= \langle \rho, \tau t^n\varphi(\hat{x}) + o(t^n)\rangle$$

$$= \tau t^n\langle \rho, w + o(1)\rangle < 0,$$

得出矛盾. 因此 $f^{[n]} \in O_{n-1}(f)$.

结合(i) 与(ii) 即推出(iii). □

注 1° 由 7.4.3, 若 $f \in C^1$, $R(f'(x_0)) = W$, $f^{[i]} = 0 (1 \leqslant i < n)$, $f^{[n]}(D, x_0, \cdots, x_{n-1}, 0)$ 存在, 则 7.6.1(iii) 之条件满足且等式(5) 成立.

2° 取 $n = 1$ 从 7.6.1 推出: 若 $f^{[1]}(D, x_0, \cdot)$ 凸且 $\exists \hat{x} \in X$: $f^{[1]}(D, x_0, \hat{x}) \leqslant 0$, 则 $F_f^1(D, x_0) = \{x : f^{[1]}(D, x_0, x) \ll 0\}$. 特别, 若 $T = f'(x_0)$ 存在且 $T^{-1}W_-^\circ \neq \varnothing$, 则 $F_f(x_0) = T^{-1}W_-^\circ$. 由此可见, 7.6.1 是 6.2.2 的推广.

7.6.2 引理 若 x_0 是问题(1) 的局部弱有效解, 则对任何 $E \subset X$ 有 $F_f^n(E) \cap K^n(E \cap M) = \varnothing$.

证 若 $z \in K^n(E \cap M)$, 则 $\exists t_k \downarrow 0, z_k \to z$, 使 $x^n(t_k, z_k) \in E \cap M$. 因 $x^n(t_k, z_k) \to x_0 (k \to \infty)$, $f(x^n(t_k, z_k)) \ll f(x_0)$ 必不满足, 因此 $z \overline{\in} F_f^n(E)$. □

注 取 $n = 1$, $E = X$ 从 7.6.2 得出 6.2.1.

7.6.3 定理 设 x_0 是问题(1) 的局部弱有效解, $K \subset F^n(D)$ 是非空开凸集; $\varphi \stackrel{\mathrm{def}}{=} f^{[n]}(x_0, \cdots, x_{n-1}, \cdot)$ 与 $\psi \stackrel{\mathrm{def}}{=} g^{[n]}(x_0, \cdots, x_{n-1}, \cdot)$ 凸且 usc, $h \in C^n$, $R(B) = Z$, $B = h'(x_0)$; 以下条件满足:

$$f^{[i]} \in K_{i-1}(f), g^{[i]} \in K_{i-1}(g), h^{[i]} = 0 \ (1 \leqslant i < n), \quad (6)$$

其中 $f^{[i]} = f^{[i]}(x_0, x_1, \cdots, x_i)$; $g^{[i]}, h^{[i]}$ 仿此. 则有 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in$

$Q_{n-1}(f)\times Q_{n-1}(g)\times Z^*$，使得

$$L^{[n]}(x_0,\cdots,x_{n-1},K)\geqslant 0, \tag{7}$$

其中 $L=\bar{\rho}f+\bar{\lambda}g+\bar{\mu}h$. 若 $\exists\hat{x}\in K:\psi(\hat{x})\in O_{n-1}(g),h^{[n]}(x_0,\cdots,x_{n-1},\hat{x})=0$,则 $\bar{\rho}\neq 0$.

证 令 $\xi=h^{[n]}(x_0,\cdots,x_{n-1},\cdot)$,则 $\xi(x)=Bx+h^{(n)}(x_n,x_1,\cdots,x_{n-1})$ 是仿射函数(§7.4(11)). 问题

$$\exists x\in K:\varphi(x)\in O_{n-1}(f),\psi(x)\in O_{n-1}(g),\xi(x)=0 \tag{8}$$

必无解. 因若 x 是(8)的解,则由条件(6)及 7.6.1,7.5.3 与 7.5.5 推出

$$x\in K\cap F_f^n(D)\cap F^n(G)\cap K^n(H)\subset F_f^n(D)\cap K^n(M)$$

(其中 $F^n(G)=F^n(G,x_0,x_1,\cdots,x_{n-1})$,$K^n(H)$ 仿此),这与 7.6.2 矛盾. 因 K 是开凸集,φ 与 ψ 凸且 usc,ξ 是开映射(1.1.7),故可应用 3.4.2 得出 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in Q_{n-1}(f)\times Q_{n-1}(g)\times Z^*$,使得

$$L^{[n]}(x_0,\cdots,x_{n-1},K)=(\bar{\rho}\varphi+\bar{\lambda}\psi+\bar{\mu}\xi)(K)\geqslant 0.$$

其次,设定理所设附加条件满足而 $\bar{\rho}=0$,则

$$0\leqslant\bar{\lambda}\psi(\hat{x})+\bar{\mu}\xi(\hat{x})=\bar{\lambda}\psi(\hat{x})\leqslant 0,$$

这与 $\psi(\hat{x})\in O_{n-1}(g)$ 一起推出 $\bar{\lambda}=0$. 于是 $0\neq\mu\in R(\xi)^*=R(B)^{\perp}=\{0\}$,得出矛盾. 可见 $\bar{\rho}\neq 0$. □

7.6.3 推广了[92,Th.5.1],[140,Th.5.1],[148,Th.5.1].

在 7.6.3 中取 $n=1$ 得到:

7.6.4 推论 设 x_0 是问题(1)的局部弱有效解,$K\subset F_D(x_0)$ 是一非空开凸集,$f^{[1]}(x_0,\cdot)$ 与 $g^{[1]}(x_0,\cdot)$ 凸且 usc,$h\in C^1$,$B=h'(x_0)$,$R(B)=Z$,则有 $0\neq(\bar{\rho},\bar{\lambda},\bar{\mu})\in W_+^*\times Q\times Z^*$,使得 $L^{[1]}(x_0,K)\geqslant 0$,此处 $L=\bar{\rho}f+\bar{\lambda}g+\bar{\mu}h$,$Q=Y_+^*\cap\{g(x_0)\}^{\perp}$. 若 $\exists\hat{x}\in K\cap N(B):g^{[1]}(x_0,\hat{x})\in Y^{\circ}+\mathbf{R}g(x_0)$,则 $\bar{\rho}\neq 0$.

容易看出,7.6.4 推广了 6.2.3,[8,Th.4.1].[8]对一阶条件有一较详细的考察.

在 7.6.3 中取 $n=2$ 得到:

7.6.5 推论　设 x_0 是问题(1)的局部弱有效解,$K \subset F^2(D, x_0, x_1)$ 是一非空开凸集,$f^{[2]}(x_0, x_1, \cdot)$ 与 $g^{[2]}(x_0, x_1, \cdot)$ 凸且 usc, $h \in C^2, B = h'(x_0), R(B) = Z; x_1 \in N(B)$ 且

$$f^{[1]}(x_0, x_1) \leqslant 0, g^{[1]}(x_0, x_1) \in Y + \mathbf{R}g(x_0). \tag{9}$$

则存在 $0 \neq (\bar{\rho}, \bar{\lambda}, \bar{\mu}) \in W_+^* \times Q \times Z^*$,使得

$$\rho f^{[1]}(x_0, x_1) = \bar{\lambda} g^{[1]}(x_0, x_1) = 0; \tag{10}$$

$$L^{[2]}(x_0, x_1, K) \geqslant 0, \tag{11}$$

其中 $L = \bar{\rho} f + \bar{\lambda} g + \bar{\mu} h$. 若 $\exists \hat{x} \in K$,使得

$$g^{[2]}(x_0, x_1, \hat{x}) \in Y_-^\circ + \text{span}\{g(x_0), g^{[1]}(x_0, x_1)\},$$

$2B\hat{x} + h''(x_0)x_1^2 = 0$,则 $\bar{\rho} \neq 0$.

对于问题

$$\min f(x),\ g(x) \leqslant 0,\ x \in D, \tag{12}$$

可以应用 3.4.8(而不是 3.4.2)得到较好的结果.

7.6.6 定理　设 x_0 是问题(12)的局部弱有效解,$K \subset K^n(D)$ 是一非空集,φ, ψ 依 7.6.3,(φ, ψ) 在 K 上为 $W_+ \times Y_+$ - 次类凸;

$$f^{[i]} \in K_{i-1}(f),\ g^{[i]} \in K_{i-1}(g), 1 \leqslant i < n. \tag{13}$$

则有 $0 \neq (\bar{\rho}, \bar{\lambda}) \in Q_{n-1}(f) \times Q_{n-1}(g)$,使得不等式(7)成立,此处 $L = \bar{\rho} f + \bar{\mu} g$. 若 $\exists \hat{x} \in K: \psi(\hat{x}) \in O_{n-1}(g)$,则 $\bar{\rho} \neq 0$.

证　若问题

$$\exists x \in K: \varphi(x) \in O_{n-1}(f), \psi(x) \in O_{n-1}(g) \tag{14}$$

有解 x,则如同 7.5.3 之证有

$$x \in K \cap F_f^n(D) \cap F^n(G) \subset F_f^n(D) \cap K^n(D \cap G),$$

得出矛盾. 因此问题(14)无解. 于是由 3.4.8 有 $0 \neq (\bar{\rho}, \bar{\lambda}) \in Q_{n-1}(f) \times Q_{n-1}(g)$,使得式(7)成立. 余下的结论是明显的. □

在 7.6.6 中取 $n = 1$ 得到:

7.6.7 推论　设 x_0 是问题(12)的局部弱有效解,$K \subset K_D(x_0)$ 是一非空集,$(f^{[1]}(x_0, \cdot), g^{[1]}(x_0, \cdot))$ 在 K 上为 $W_+ \times Y_+$ - 次类凸,则有 $0 \neq (\bar{\rho}, \bar{\lambda}) \in W_+^* \times Q$,使得 $L^{[1]}(x_0, K) \geqslant 0$,此处 L

$= \rho f + \lambda g$. 若 $\exists \hat{x} \in K$: $g^{[1]}(x_0, \hat{x}) \in Y_-^\circ + \mathbf{R} g(x_0)$，则 $\bar{\rho} \neq 0$.

在 7.6.6 中取 $n = 2$ 得到：

7.6.8 推论 设 x_0 是问题(12)的局部弱有效解，$K \subset K^2(D, x_0, x_1)$ 是一非空集，$(f^{[2]}(x_0, x_1, \cdot), g^{[2]}(x_0, x_1, \cdot))$ 在 K 上为 $W_+ \times Y_+$ 次类凸；条件(9)满足. 则有 $0 \neq (\bar{\rho}, \bar{\lambda}) \in W_+^* \times Q$，使得(10)，(11)成立(但取 $L = \bar{\rho} f + \bar{\lambda} g$). 若 $\exists \hat{x} \in K$: $g^{[2]}(x_0, x_1, \hat{x}) \in Y_-^\circ + \mathbf{R} g(x_0) + \mathbf{R} g^{[1]}(x_0, x_1)$，则 $\bar{\rho} \neq 0$.

参考文献：[8,38,70,77,92,140,148].

第八章　选择论题

本章挑选的几个论题在最优化理论中都已显示出无可置疑的重要性.不过,本书已来不及作深入的展开,此处所选择的材料仅带有导引性质.

§8.1　具多值约束函数的极小问题

本节考虑以下 VOP

$$\min f(x),\ a \in Gx,\ x \in D, \tag{1}$$

其中 $f: X \to W, a \in Y, G: X \to 2^Y, D \subset X$. 假定 W 中已由闭凸锥 W_+ 导入序 $\leqslant$. 问题(1) 具有很大的一般性. 例如,若以 $Y \times Z$ 替代 Y,取 $a = 0, Gx = (g(x) + Y_+) \times \{h(x)\}$,则问题(1) 成为

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D. \tag{2}$$

若取 $a = 0, Gx = g(x) + Y_+$,则问题(1) 就是

$$\min f(x),\ g(x) \leqslant 0,\ x \in D. \tag{3}$$

可见本书前面所讨论的最优化问题全是问题(1)的特例.

问题(1)的可行集是 $M = D \cap G^{-1}(a)$. 仿照第六章的作法,定义(1) 的一个 Lagrange 函数:

$$L(x, \rho, \lambda) = \rho f(x) + \lambda(Gx - a), \tag{4}$$

它是一个定义于 $(D \cap D_G) \times W_+^* \times Y^*$ 上的多值函数.

择一定理是研究问题(1) 的基本工具之一. 设 $\bar{x} \in M$,令 $\tilde{f}(x) = f(x) - f(\bar{x})$. $\bar{x}$ 是问题(1) 的弱有效解意味着问题

$$\exists x \in D:\ \tilde{f}(X) \ll 0,\ 0 \in Gx - a \tag{5}$$

无解. 问题(5)关联着备择的命题:

$$\exists(\rho,\lambda)\in(W_+^*\backslash\{0\})\times Y^*:L(D\cap D_G,\rho,\lambda)\geqslant\rho f(\bar{x}).\quad(6)$$

若 (ρ,λ) 满足问题(6)，则以 $x=\bar{x}$ 代入 $L(x,\rho,\lambda)\geqslant\rho f(\bar{x})$ 得 $\lambda(G\bar{x}-a)\geqslant 0$，即 $\lambda\in(G\bar{x}-a)^*$. 例如，令 $a=0,Gx=g(x)+Y_+$ 得

$$\lambda\in(g(\bar{x})+Y_+)^*=Y_+^*\cap\{g(\bar{x})\}^\perp=Q.$$

基于以上分析，从 3.4.6 得到：

8.1.1 定理 设 $\Phi=(f,G)$. 假定：(i) 任给 $t\in J$ 与 W 的 0-邻域 V，有 $tR(\Phi)+t'R(\Phi)\subset R(\Phi)+W_+\times 0+V\times 0$；(ii) $\exists x_0\in D$：f 在 x_0 为 usc；(iii) $\exists y_0\in Gx_0$，任给 X 的 0- 邻域 N，存在 Y 的 0- 邻域 V：$y_0+V\subset G(D\cap(x_0+N))$；(iv) $(R(G)-a)^*=\{0\}$. 则 $\bar{x}\in M$ 是问题(1) 的弱有效解的充要条件是问题(6) 有解 $(\rho,\lambda)\in(W_+^*\backslash\{0\})\times Y^*$，其中 $\lambda\in((G\bar{x}-a)^*$.

若取 $a=0,Gx=g(x)+Y_+$，则 (f,g) 在 D 上为 $W_+\times Y_+$- 类凸 $\Rightarrow(f,G)$ 在 D 上为 $W_+\times 0$- 类凸 $\Rightarrow$ 8.1.1 中条件(i) 满足；$R(G)^*=(g(D)+Y_+)^*=g(D)^*\cap Y_+^*$；$\rho f(x)+\lambda G(x)\geqslant\rho f(\bar{x})\Leftrightarrow\rho f(x)+\lambda g(x)\geqslant\rho f(\bar{x})$ 且 $\lambda\in Y_+^*$. 利用这些事实，可从 8.1.1 推出关于问题(3) 的适当结论.

下面设 f 为实值函数，$\alpha=\inf f(M)$. 称

$$\max\inf L(D,\lambda)=\beta,\ \lambda\in Y^*\quad(7)$$

为问题(1)的对偶问题. 注意恒有 $\beta\leqslant\alpha$：

$$\begin{aligned}\beta&=\sup_{\lambda\in Y^*}\inf L(D,\lambda)\\&=\sup_{\lambda\in Y^*}\inf_{x\in D}\inf_{y\in Gx}[f(x)+\lambda(y-a)]\\&\leqslant\inf_{x\in D}\inf_{y\in Gx}[f(x)+\sup_{\lambda\in Y^*}\lambda(y-a)]\\&=\inf_{x\in D}\inf_{y=a\in Gx}f(x)=\alpha.\end{aligned}$$

现在建立一个类似于 5.2.9 的结果.

8.1.2 定理 设 $D\subset D_G,\Phi=(f,G)$ 为 $\mathbf{R}_+\times 0$- 次类凸，α 有限，且以下条件之一满足：(i) 存在 $\sigma>\alpha$ 与 Y 的 0- 邻域 V，使 $C\cap((-\infty,\sigma]\times V)$ 闭，$C=R(\Phi)+\mathbf{R}_+\times\{-a\}$；(ii) $\Omega\overset{\text{def}}{=}\{(x,r,$

$y): x \in D, f(x) \leqslant \alpha + r, y \in Gx - a\}$ 闭，存在 $\sigma > \alpha$ 与 Y 的闭 0- 邻域 V，使 $f^{-1}((-\infty, \sigma]) \cap G^{-1}(V)$ 为紧集；(iii) X 为自反 B-空间，当 X 中用弱拓扑时 Ω 为闭集，且有 $(\rho, \lambda) \in \mathbf{R}_+ \times Y^*$，使得

$$\liminf_{y \in Gx, \|x\| \to \infty} \rho f(x) + \lambda(y - a) > \rho\alpha. \tag{8}$$

则问题(1)有最优解 $\bar{x}$；$\forall \varepsilon > 0, \exists \lambda_\varepsilon \in Y^*$，使得

$$\alpha \leqslant \inf L(D, \lambda_\varepsilon) + \varepsilon, \tag{9}$$

因而 $\alpha = \beta$.

证 不妨设条件(i)满足. 取 $x_n \in M$，使 $f(x_n) < \alpha + n^{-1}$. 可设 $\alpha + n^{-1} < \sigma$，于是

$$(\alpha + n^{-1}, 0) \in C \cap ((-\infty, \sigma] \times V).$$

由条件(i)有 $(\alpha, 0) \in C$，故有 $\bar{x} \in D: f(\bar{x}) \leqslant \alpha, 0 \in G\bar{x} - a$，可见 $\bar{x}$ 是(1) 的最优解.

因问题

$$\exists x \in D: f(x) - \alpha < 0, \ 0 \in Gx - a$$

无解，故由条件(i)与 3.4.9，$\forall \varepsilon > 0, \exists \lambda_\varepsilon \in Y^*$，使得 $L(D, \lambda_\varepsilon) > \alpha - \varepsilon$，这推出(9) 成立. 由(9) 得出 $\alpha \leqslant \beta + \varepsilon$；令 $\varepsilon \downarrow 0$ 得 $\alpha = \beta$. □

8.1.3 定义 若 $\mathrm{Gr}\, G$ 为凸锥且 $0 \in G(0)$，则称 G 为**凸过程**；若再假定 $\mathrm{Gr}\, G$ 为闭集，则称 G 为**闭凸过程**.

G 是凸过程的充要条件是 $D_G \neq \varnothing$ 且

$$\begin{cases} tGx + t'Gy \subset G(tx + t'y), \\ \alpha Gx \subset G(\alpha x) \ (t \in J, \alpha \geqslant 0). \end{cases} \tag{10}$$

因此，线性映射必为凸过程，而凸过程必为凸映射. 若 $Gx = g(x) + Y_+$，则 G 是凸过程 $\Leftrightarrow g$ 为 Y_+ - 凸映射且 $\alpha g(x) \geqslant g(\alpha x)$ ($\forall \alpha \geqslant 0$)；若 g 关于 Y_+ 为 $*$ lsc，则 $\mathrm{Gr}\, G$ 是闭的.

具凸过程约束的最优化问题曾为一些学者所研究(参看[18, 83]). 下面结果的基本思想属于[83].

8.1.4 定理([107]) 设 G 为闭凸过程，f 是包含 $M \overset{\text{def}}{=} G^{-1}(a)$

的某开集内的 C^{1-0} 实函数，且在 $\bar{x} \in M$ 关于 M 为广义伪凸；$D = X, T(x,\alpha) = Gx - \alpha a$，$(\mathrm{Gr}\, T)^* + 0 \times 0 \times Y^*$ 弱*闭．则 $\bar{x}$ 是问题(1)的最优解的充要条件是，存在 $u \in \partial f(\bar{x})$，使得

$$\exists \lambda \in Y^*: u \in -G^*\lambda,\ u(\bar{x}) = -\lambda(a). \qquad (11)$$

证 因 G 为凸过程，故 M 是凸集，从而 $\bar{x}$ 是问题(1)的最优解的充要条件是存在 $u \in \partial f(\bar{x}) \cap (M - \bar{x})^*$ (2.2.9).

1° 证 $u \in (M - \bar{x})^* \Leftrightarrow (u, -u(\bar{x})) \in [T^{-1}(0)]^*$．任给 $u \in (M - \bar{x})^*$，$(x,\alpha) \in T^{-1}(0)$，有 $0 \in Gx - \alpha a$．取充分小的 $\varepsilon > 0$．由 $\alpha\varepsilon a \in G(\varepsilon x)$ 与 $a \in G\bar{x}$ 及(10)得

$$(1 + \alpha\varepsilon)a \in G\bar{x} + G(\varepsilon x) \subset G(\bar{x} + \varepsilon x).$$

可设 $1 + \alpha\varepsilon > 0$，于是 $a \in G((\bar{x} + \varepsilon x)/(1 + \alpha\varepsilon))$，从而

$$\frac{\bar{x} + \varepsilon x}{1 + \alpha\varepsilon} - \bar{x} \in M - x;$$

$$0 \leqslant u\left(\frac{\bar{x} + \varepsilon x}{1 + \alpha\varepsilon} - \bar{x}\right) = \frac{\varepsilon[u(x) - \alpha u(\bar{x})]}{1 + \alpha\varepsilon},$$

故得 $u(x) - \alpha u(\bar{x}) \geqslant 0$，即 $(u, -u(\bar{x})) \in [T^{-1}(0)]^*$．反之，若 $(u, -u(\bar{x})) \in [T^{-1}(0)]^*$，$x \in M$，则 $0 = a - a \in T(x,1)$，于是 $u(x) - u(\bar{x}) \geqslant 0$，可见 $u \in (M - \bar{x})^*$.

2° 证 $-[T^{-1}(0)]^* = T^*Y^*$．由 $\mathrm{Gr}\, G$ 为闭凸锥推出 $\mathrm{Gr}\, T$ 亦为闭凸锥．因 $(\mathrm{Gr}\, T)^* + 0 \times 0 \times Y^*$ 弱*闭，故由 3.1.1 得出所要等式.

3° 证 $T^*Y^* = \{(u, -\lambda(a)) : u \in G^*\lambda, \lambda \in Y^*\}$．设 $\lambda \in Y^*$，$(u,k) \in T^*\lambda$．任取 $(x,y) \in \mathrm{Gr}\, G$，$t \in \mathbf{R}$，有 $y - ta \in T(x,t)$，$(x, t, y - ta) \in \mathrm{Gr}\, T$，于是

$$\begin{aligned} 0 &\leqslant \langle(-u, -k, \lambda), (x, t, y - ta)\rangle \\ &= \lambda(y) - u(x) - t[\lambda(a) + k], \end{aligned}$$

即 $t[\lambda(a) + k] \leqslant \lambda(y) - u(x)$．因此不等式在 $t \to \pm\infty$ 时皆成立，故必 $\lambda(a) + k = 0$，从而 $\lambda(y) - u(x) \geqslant 0$．这表明 $k = -\lambda(a)$，$u \in G^*\lambda$．反之，如果 $\lambda \in Y^*$，$u \in G^*\lambda$，则对任给 $(x, \alpha, y - \alpha a) \in$

$\mathrm{Gr}\,T$(这意味着 $y \in Gx$)有

$$\langle(-u,\lambda(a),\lambda),(x,\alpha,y-\alpha a)\rangle$$
$$=\lambda(y)-u(x)=\langle(-u,\lambda),(x,y)\rangle \geqslant 0,$$

可见 $(u,-\lambda(a)) \in T^*\lambda$.

综合以上所证,得出定理结论. □

取 $Gx = g(x) + Y_+, a = 0$ 从 8.1.4 得:

8.1.5 推论 设 g 为凸映射且 $\alpha g(x) \geqslant g(\alpha x)$ $(\forall \alpha \geqslant 0)$;$g$ 关于 Y_+ 为 $*$ lsc;f 在包含 $M \stackrel{\mathrm{def}}{=} g^{-1}(Y_-)$ 的某个开集内为 C^{1-0} 实函数,且在 $\bar{x} \in M$ 关于 M 为广义伪凸;$D = X$;$\varphi(x,\alpha) = g(x) - \alpha a$,$(\mathrm{epi}\,\varphi)^* + 0 \times 0 \times Y^*$ 为弱*闭集. 则 $\bar{x}$ 是问题(3)的最优解 $\Leftrightarrow$ 存在 $u \in \partial f(\bar{x})$,使得对某个 $\lambda \in Y^*$ 有 $(u,\lambda) \in (\mathrm{epi}\,g)^*$ 且 $u(\bar{x}) = 0$.

下面给出一个基于 5.6.6 的结果([109]).

8.1.6 定理 假定:(i) $f: X \to \overline{\mathbf{R}}$ 凸、lsc 且 $D_f \neq \varnothing$;(ii) G 为闭凸过程;(iii) $0 \in \mathrm{sqri}\,G(D_f)$ 且 $0 \in \mathrm{sqri}\,(D_f - D_G)$. 设 $a = 0$, $D = X$, $\alpha = \inf\{f(x): 0 \in Gx\}$ 有限,则问题

$$\max\,[-f^*(-u)] = \beta,\ u \in R(G^*)$$

有解且 $\alpha = \beta$.

证 令 $\varphi(x,y) = f(x) + \delta_{\mathrm{Gr}G}(x,y)$,则 φ 凸且为 lsc, $D_\varphi = (D_f \times Y) \cap \mathrm{Gr}\,G \neq \varnothing$. 由条件 (iii)有

$$0 \in \mathrm{sqri}\,G(D_f) = \mathrm{sqri}\,(\pi D_\varphi),$$

其中 $\pi: X \times Y \to Y$ 是投影. 于是由 5.6.6 有 $\lambda \in Y^*$,使得 $\alpha = -\varphi^*(0,\lambda)$. 因此

$$\alpha = \inf\{\varphi(x,y) - \lambda(y) : (x,y) \in X \times Y\}$$
$$= \inf\{f(x) + \delta(x,y) - \lambda(y) : (x,y) \in X \times Y\},$$

其中 $\delta = \delta_{\mathrm{Gr}G}$. 由 $0 \in \mathrm{sqri}\,(D_f - D_G)$ 与

$$(D_G - D_f) \times Y = \mathrm{Gr}\,G - D_f \times Y$$

推出 $0 \in \mathrm{sqri}(D_f \times Y - D_\delta)$. 对 f 与 $\delta(x,y) - \lambda(y)$ 应用 5.6.5 得

$$\alpha = \max\{-f^*(-u):\ u \in G^*(-\lambda)\}$$
$$= \max\{-f^*(-u):\ u \in R(G^*)\}. \qquad \square$$

参考文献：[18,83,107,109].

§8.2 具无限个不等式约束的极小问题

本节考虑如下最优化问题：

$$\min f(x) = \alpha, \varphi(x,y) \leqslant 0 (\forall y \in K), x \in D, \tag{1}$$

其中 $f: X \to \overline{\mathbf{R}}, f|D \not\equiv \infty, \varphi: D \times K \to \mathbf{R}$. 问题(1)有很大的一般性. 例如，问题

$$\min f(x), g(x) \leqslant 0, h(x) = 0, x \in D \tag{2}$$

与
$$\min f(x), g(x) \leqslant 0, x \in D \tag{3}$$

都可写成(1)的形式，只需分别取 $\varphi(\cdot,\lambda,\mu) = \lambda g + \mu h, K = Y_+^* \times Z^*$ 与 $\varphi(\cdot,\lambda) = \lambda g, K = Y_+^*$，此处 $g: D \to Y, h: D \to Z$. 当 $\dim X < \infty$，K 为无限集时，(1) 称为**半无限最优化问题**.

对于问题(1)有两种定义 Lagrange 函数的方法，它们所据的思路及所需的条件互不相同，且导致不同的对偶问题. 首先设 K 是某个 LCS 中的锥，$\varphi(x,\cdot)$ 为(1次)正齐次函数，定义问题(1)的 Lagrange 函数为

$$L(x,y) = f(x) + \varphi(x,y), (x,y) \in D \times K. \tag{4}$$

以 M 记问题(1)(及(2),(3))的可行集. 易验证

$$\sup L(x,K) = \begin{cases} f(x), & x \in M, \\ \infty, & \text{否则}; \end{cases}$$

$$\inf_{x\in D} \sup L(x,K) = \inf_{x\in M} f(x) = \alpha.$$

于是，(1)有如下对偶问题(参考§5.1)：

$$\max \inf L(D,y) = \beta,\ y \in K. \tag{5}$$

若取 $\varphi(\cdot,\lambda) = \lambda g$，则 $L(x,\lambda) = f(x) + \varphi(x,\lambda)$ 就是问题(3)的通常的 Lagrange 函数，而与(5)相当的对偶问题就是通常的

Lagrange 对偶(参看 §5.2(5*)).

今用 §3.6 与 §5.1 的定理来得出关于问题(1) 的最优性条件与对偶结果. 设 $\bar{x} \in M$, 令 $\tilde{f}(x) = f(x) - f(\bar{x})$, $\hat{f}(x) = f(x) - \alpha$, 后者自然要求 α 有限. 问题

$$\exists x \in D: \hat{f}(x) < 0,\ \varphi(x,K) \leqslant 0 \tag{6}$$

必无解; 而 $\bar{x}$ 是问题(1) 的最优解相当于问题

$$\exists x \in D: \tilde{f}(x) < 0,\ \varphi(x,K) \leqslant 0 \tag{7}$$

无解. 形成备择的命题:

$$\exists y \in K: \hat{f}(x) + \varphi(x,y) \geqslant 0 (\forall x \in D); \tag{8}$$

$$\exists y \in K: \tilde{f}(x) + \varphi(x,y) \geqslant 0 (\forall x \in D). \tag{9}$$

若式(6)与式(8)两择一, 则式(8)必有解 $y = \bar{y}$, 于是 $L(D,\bar{y}) \geqslant \alpha$. 这推出 $\alpha \leqslant \inf L(D,\bar{y}) \leqslant \beta$, 从而 $\alpha = \inf L(D,\bar{y}) = \beta$, 即 $\bar{y}$ 是问题(5) 的解且 $\alpha = \beta$. 若式(7) 与式(9) 两择一, 则 $\bar{x}$ 是问题(1) 的最优解的充要条件是: 式(9) 有解 $y = \bar{y}$, 这相当于

$$f(\bar{x}) = \min L(D,\bar{y}),\ \varphi(\bar{x},\bar{y}) = 0. \tag{10}$$

由 5.1.1, (10) $\Leftrightarrow$ $(\bar{x},\bar{y})$ 是 $L(x,y)$ 的鞍点 $\Leftrightarrow$ $\bar{x}$ 与 $\bar{y}$ 分别为式(1) 与式(5) 的解且 $\alpha = \beta$. 对于问题(2) 与式(3), 式(8) 正好归结为 Lagrange 函数的鞍点条件(参看 §5.2(2), (6)).

以上分析表明, 一旦指明了式(6) 与式(8) 或式(7) 与式(9) 的择一性, 就可获得较强的对偶性结果或最优性条件. 所需的择一性依赖于 3.6.2.

8.2.1 定理 假定: (i) 锥 K 有紧基 B; (ii) φ 对 y 正齐次、usc 且在 B 上次类凹; (iii) $(f,\varphi(\cdot,y))(y \in B)$ 近次类凸; (iv) $\{y \in K: \varphi(D,y) \geqslant 0\} = \{0\}$. 若 α 有限, 则问题(5) 有解且 $\alpha = \beta$. $x \in M$ 是问题(1) 的解 $\Leftrightarrow$ 问题(8) 有解 $\bar{y} \in K$ $\Leftrightarrow$ $\exists \bar{y} \in K$, 使 $(\bar{x},\bar{y})$ 是 $L(x,y)$ 的鞍点.

证 令 $Q = \{(t,t'y): t \in J, y \in B\}$, 则 Q 是 $\mathbf{R}_+ \times K$ 的紧基(参考 3.6.6 之证). 任给 $x \in D$, $v = (t,t'y) \in Q$, 令

$$\Gamma(x,v) = t[f(x) - r] + \varphi(x, t'y),$$

$r \in \mathbf{R}$ 是任意取定的. 然后依 $\Gamma(x,\rho v) = \rho\Gamma(x,v)$ 将 Γ 扩张到 $D \times \mathbf{R}_+ \times K$ 上. $\Gamma(x,\cdot)$ 显然为正齐次且 usc. 由条件(iii) 推出 $\Gamma(x,v)$ 对 x 在 D 上近次类凸;由条件(ii) 推出 $\Gamma(x,v)$ 对 v 在 Q 上次类凹.

设 α 有限,取 $r = \alpha$,则式(6) 无解,这推出问题

$$\exists x \in D: \Gamma(x,Q) < 0 \tag{11}$$

无解. 否则,有 $x \in D, \forall (t,y) \in J \times B$,有

$$t[f(x) - \alpha] + \varphi(x,t'y) < 0.$$

取 $t = 0$ 得 $\varphi(x,B) < 0$,从而 $\varphi(x,K) \leqslant 0$;取 $t - 1$ 得 $\hat{f}(x) < 0$,因而 x 是式(6) 的解,得出矛盾. 因为式(11) 无解,故由 3.6.2 有 $v = (t,t'y) \in Q$,使 $\Gamma(D,v) \geqslant 0$. 若 $t = 0$,则 $\Gamma(D,v) = \varphi(D,y) \geqslant 0$,于是由条件(iv) 有 $y = 0$,与 $y \in B$ 矛盾. 故 $t > 0$. $\forall x \in D$, 有

$$0 \leqslant \hat{f}(x) + t^{-1}\varphi(x,t'y) = \hat{f}(x) + \varphi(x,\bar{y}),$$

其中 $\bar{y} = t^{-1}t'y \in K$. 这表明式(8) 成立.

同理,式(7) 与式(9) 必两择一. □

类似的结果参看[43,Th. 3.1].

取 $\varphi(x,\lambda) = \lambda g(x)$ 从 8.2.1 得出:

8.2.2 推论 设 Y_+^* 有弱*紧凸基;(f,g) 在 D 上 $\mathbf{R}_+ \times Y_+$-次类凸;$Y_+^* \cap g(D)^* = \{0\}$. 若 $\alpha \overset{\text{def}}{=} \inf\limits_{x \in M} f(x)$ 有限,则问题(3) 的对偶问题

$$\max \inf\ (f + \lambda g)(D) = \beta,\ \lambda \in Y_+^*$$

有解且 $\alpha = \beta$. $\bar{x} \in M$ 是问题(3) 的最优解 $\Leftrightarrow \exists \bar{\lambda} \in Q \overset{\text{def}}{=} Y_+^* \cap \{g(\bar{x})\}^-$; $f(\bar{x}) = \min\ (f + \bar{\lambda} g)(D) \Leftrightarrow \exists \bar{\lambda} \in Q$: $(\bar{x},\bar{\lambda})$ 是 $L(x,\lambda) = f(x) + \lambda g(x)$ 在 $D \times Y_+^*$ 上的鞍点.

8.2.2 包含了 4.4.3 与 5.2.8 的部分结论.

若 D 凸且 $\varphi(\cdot,y)$ 凸,则 M 凸. 若再假定 f 在 $\bar{x}$ 关于 M 为广义伪凸,则 $\bar{x}$ 是问题(1) 的最优解 $\Leftrightarrow \partial f(\bar{x}) \cap (M - \bar{x})^* \neq \varnothing$

(2.2.9). 注意到 $u \in (M - \bar{x})^*$ 相当于 $\bar{x}$ 是问题

$$\min u(x), \varphi(x,y) \leqslant 0 (\forall y \in K), x \in D \tag{12}$$

的最优解，将 8.2.1 用于问题(12)得出：

8.2.3 推论 设 D 凸，锥 K 有紧凸基 B；f 在 $\bar{x} \in M$ 关于 M 为广义伪凸；φ 在 $D \times B$ 上为鞍函数，$\varphi(x,\cdot)$ 正齐次且 usc；$\{y \in K : \varphi(D,y) \geqslant 0\} = \{0\}$. 则 $\bar{x}$ 是问题(1) 的最优解的充要条件是：存在 $u \in \partial f(\bar{x}), \bar{y} \in K$，使得

$$u(\bar{x}) = \min_{x \in D}[u(x) + \varphi(x,\bar{y})], \ \varphi(\bar{x},\bar{y}) = 0. \tag{13}$$

若 $D = X$，则条件(13) 推出

$$-[u(x) - u(\bar{x})] \leqslant \varphi(x,\bar{y}) - \varphi(\bar{x},\bar{y}), \forall x \in X,$$

这意味着 $-u \in \partial_x \varphi(\bar{x},\bar{y})$. 因此，当 $D = X$ 时 8.2.3 的结论可表成：$\bar{x}$ 是问题(1) 的最优解的充要条件是：存在 $\bar{y} \in K$，使得

$$0 \in \partial f(\bar{x}) + \partial_x \varphi(\bar{x},\bar{y}), \varphi(\bar{x},\bar{y}) = 0. \tag{14}$$

式(14)可看作是一非光滑的"Lagrange 条件".

将 8.2.3 用到问题(3)得：

8.2.4 推论 设 D 凸，Y_+^* 有弱*紧凸基；f 在 $\bar{x} \in M$ 关于 M 为广义伪凸；g 凸且 $Y_+^* \cap g(D)^* = \{0\}$. 则 $\bar{x}$ 是问题(3) 的最优解的充要条件是：存在 $u \in \partial f(\bar{x})$ 与 $\bar{\lambda} \in Q$，使得

$$u(\bar{x}) \leqslant u(x) + \bar{\lambda} g(x) (\forall x \in D).$$

若 $D = X$，则上述结论可表成：存在 $\bar{\lambda} \in Q$，使得

$$0 \in \partial f(\bar{x}) + \partial(\bar{\lambda} g)(\bar{x}).$$

现在将 8.2.1 用于所谓齐次规划问题：

$$\min f(x) = \alpha, \varphi(x,y) \geqslant F(y) (\forall y \in K), x \in D; \tag{15}$$

$$\max F(y) = \beta, \varphi(x,y) \geqslant f(x) (\forall x \in D), y \in K, \tag{16}$$

其中 D,K 是锥，$f(x), F(y)$ 与 $\varphi(x,y)$（对每个变元）是正齐次函数. 分别以 M 与 N 记式(15) 与(16) 的可行集. 令 $L(x,y) = f(x) + F(y) - \varphi(x,y)$，则易验证

$$\inf L(D,y) = \begin{cases} F(y), & y \in N, \\ -\infty, & \text{否则}, \end{cases}$$

可见(15)与(16)互为对偶. 由 8.2.1 有：

8.2.5 推论 假定：(i) K 有紧凸基；(ii) $\varphi(x,\cdot)$ 凸且为 lsc，F 凹且为 usc；(iii) $(f,-\varphi(\cdot,y))$ 近次类凸；(iv) $\{y\in K:\varphi(x,y)\leqslant F(y)(\forall x\in D)\}=\{0\}$. 若 α 有限，则问题(16) 有解且 $\alpha=\beta$. $\bar{x}\in M$ 是问题(15) 的最优解的充要条件是：存在 $\bar{y}\in K$，使得

$$f(\bar{x})=\varphi(\bar{x},\bar{y})=F(\bar{y}). \tag{17}$$

式(17)推出 $(\bar{x},\bar{y})$ 是 $L(x,y)$ 的鞍点且 $\alpha=\beta$.

在[43,80] 中可看到类似的结果.

问题(15),(16) 的一种特殊情形是：

$$\min f(x)=\alpha,\langle\lambda,g(x)\rangle\geqslant F(\lambda)(\forall\lambda\in Y_+^*),x\in D; \tag{18}$$

$$\max F(\lambda)=\beta,\langle\lambda,g(x)\rangle\leqslant f(x)(\forall x\in D),\lambda\in Y_+^*, \tag{19}$$

其中 $f(x),g(x),F(\lambda)$ 都是正齐次的. 将 8.2.5 用于问题(18),(19) 得：

8.2.6 推论 假定：(i) Y_+^* 有弱*紧凸基；(ii) F 凹且为 usc；(iii) $(f,-g)$ 在 D 上次类凸；(iv) $\{\lambda\in Y_+^*:\langle\lambda,g(x)\rangle\leqslant F(\lambda)(\forall x\in D)\}=\{0\}$. 若 α 有限，则问题(19) 有解且 $\alpha=\beta$. 问题(18) 的可行点 $\bar{x}$ 是最优解的充要条件是：存在 $\bar{\lambda}\in Y_+^*$，使得

$$f(\bar{x})=\langle\bar{\lambda},g(\bar{x})\rangle=F(\bar{\lambda}).$$

现在再回到问题(1)，但采用新的 Lagrange 函数

$$L(x,\lambda)=f(x)+\langle\lambda,\varphi(x,\cdot)\rangle,x\in D,\lambda\in F_+^*, \tag{20}$$

记号 $F=\mathbf{R}^K$ 及 F_+^* 等依 §3.5. 易验证

$$\sup L(x,F_+^*)=\begin{cases}f(x), & x\in M,\\ \infty, & \text{否则}.\end{cases}$$

这就得到问题(1)的又一个对偶(对照(5))：

$$\max\inf L(D,\lambda)=\beta,\ \lambda\in F_+^*. \tag{21}$$

现在利用 3.5.11 证明：

8.2.7 定理 假定：(i) (f,Φ) 为 $\mathbf{R}_+\times F_+$ 次类凸，Φ 依 §3.5(6)；(ii) 存在 $\sigma\in\mathbf{R}$ 与 F 的 0-邻域 V，使 $C\cap((-\infty,\sigma]\times V)$ 为闭集，C 依 §3.5(14). 若 $\sigma\geqslant\alpha>-\infty$，则 $\alpha=\beta$. 若 $\bar{x}$ 是

问题(1)的最优解,$\sigma \geqslant f(\bar{x})$,则$\forall \varepsilon > 0, \exists \lambda_\varepsilon \in F_+^*$:$(\bar{x}, \lambda_\varepsilon)$是$L(x, \lambda)$的$\varepsilon$-鞍点.若$\sigma > \alpha > -\infty$,则(1)必存在最优解.条件(ii)可代以如下条件之一:(iii) D为Hausdorff空间,f与φ对x为lsc;$\exists \sigma \in \mathbf{R}, \omega \subset K, \varepsilon > 0$,使$\{x \in D: f(x) \leqslant \sigma, \varphi(x, \omega) \leqslant \varepsilon\}$为紧集;(iv) X为自反B-空间,$D \subset X$弱闭,f与φ对x为wlsc,存在$\sigma \in \mathbf{R}$与$(\rho, \lambda) \in \mathbf{R}_+ \times F_+^*$,使得

$$\liminf_{x \in D, |x| \to \infty} \rho[f(x) - \sigma] + \langle \lambda, \varphi(x, \cdot) \rangle > 0. \qquad (22)$$

证 1° 设$\sigma \geqslant \alpha > -\infty$.因问题

$$\exists x \in D: f(x) - \alpha < 0, \ \varphi(x, K) \leqslant 0$$

无解,故由条件(i),(ii)知可用3.5.11得出:$\forall \varepsilon > 0, \exists \lambda_\varepsilon \in F_+^*$,使得(注意$\alpha$有限$\Rightarrow M \neq \varnothing$!)

$$f(x) - \alpha + \langle \lambda_\varepsilon, \varphi(x, \cdot) \rangle > -\varepsilon \quad (\forall x \in D).$$

由此推出$\alpha < L(D, \lambda_\varepsilon) + \varepsilon$,因而$\alpha \leqslant \inf L(D, \lambda_\varepsilon) + \varepsilon \leqslant \beta + \varepsilon$,令$\varepsilon \downarrow 0$,得$\alpha = \beta$.

2° 设$\bar{x}$是(1)的最优解且$\sigma \geqslant f(\bar{x})$,则以$\alpha = f(\bar{x})$代入上段证明得出:$\forall \varepsilon > 0, \exists \lambda_\varepsilon \in F_+^*$,使得

$$\sup L(\bar{x}, F_+^*) = \alpha \leqslant \inf L(D, \lambda_\varepsilon) + \varepsilon,$$

这表明$(\bar{x}, \lambda_\varepsilon)$是$L(x, \lambda)$的$\varepsilon$-鞍点.

3° 设$\sigma > \alpha > -\infty$.$\forall r \in (\alpha, \sigma], \exists x \in M: f(x) < r$,这表明$(r, 0) \in C$.因此$(\alpha, \sigma] \times 0 \subset C \cap ((-\infty, \sigma] \times V)$,于是由条件(ii)得$(\alpha, 0) \in C$,即有$\bar{x} \in D$,使$f(\bar{x}) \leqslant \alpha, \varphi(\bar{x}, \cdot) \leqslant 0$.这正表明$\bar{x}$是问题(1)的最优解.

4° 用§3.5中的方法易验知(iii)$\Rightarrow$(ii),(iv)$\Rightarrow$(ii). □

8.2.7推广了[115,Th.2.1].

注 (22)可代以较易验证的条件

$$\liminf_{x \in D, x \to \infty} f(x) > \sigma$$

或

$$\exists y \in K: \liminf_{x \in D, |x| \to \infty} \varphi(x, y) > 0.$$

8.2.8 例([115]) 考虑半无限最优化问题

$$\begin{cases}\min\ (x_1^2+x_2)=\alpha,\ x=(x_1,x_2)\in\mathbf{R}^2,\\ x_1\leqslant 0,\ -x_2\leqslant 1, n^{-1}x_1\leqslant x_2, n=3,4,\cdots.\end{cases}\tag{23}$$

这相当于在(1)中取 $D=X=\mathbf{R}^2, K=\mathbf{N}, f(x)=x_1^2+x_2$,

$$\varphi(x,n)=\begin{cases}x_1, & n=1,\\ -x_2-1, & n=2,\\ n^{-1}x_1-x_2, & n\geqslant 3.\end{cases}$$

令 $F=\mathbf{R}^{\mathbf{N}}$. 依问题(20), 问题(23)的 Lagrange 函数为

$$L(x,\lambda)=x_1^2+x_2+\lambda_1x_1-\lambda_2(x_2+1)+\sum_{n\geqslant 3}\lambda_n\left(\frac{x_1}{n}-x_2\right),$$

其中 $\lambda=(\lambda_n)\in F_+^*$. 问题(23) 的对偶为

$$\max\inf L(\mathbf{R}^2,\lambda)=\beta,\ \lambda\in F_+^*.\tag{24}$$

取定 $n\geqslant 3$,取 $\lambda_n=1$,而 $\lambda_i=0(i\neq n)$,则

$$\inf L(\mathbf{R}^2,\lambda)=\inf_{x\in\mathbf{R}^2}\left(x_1^2+\frac{x_1}{n}\right)=-\frac{1}{4n^2}.$$

这推出 $\beta\geqslant 0$. 另一方面,由 $f(0)=0,\varphi(0,\mathbf{N})\leqslant 0$ 推出 $\alpha\leqslant 0$,因此 $\alpha=\beta=0$,且 $x=0$ 是问题(23) 的解.

注意,$f(x)$ 与 $\varphi(x,n)$ 都是 x 的连续凸函数;若取 $\sigma=\varepsilon=1>0,\omega=\{2\}\subset\mathbf{N}$,则

$$\begin{aligned}&\{x\in\mathbf{R}^2: f(x)\leqslant\sigma,\varphi(x,\omega)\leqslant\varepsilon\}\\ &=\{x\in\mathbf{R}^2: -2\leqslant x_2\leqslant 1-x_1^2\}\end{aligned}$$

是紧集. 因此,可用 8.2.7 推出问题(23)有解且 $\alpha=\beta$.

参考文献: [17, 43, 80, 82, 87, 111, 113, 115, 118, 136, 140, 187].

§8.3 值 函 数

给定 $\varphi: X\times Y\to\overline{\mathbf{R}}$,如 §5.4 中所指出的,$\varphi$关联着一对最优化问题:

(P):　　$\min \varphi(x,0) = \alpha, x \in X$;

(P*):　　$\max[-\varphi^*(0,\lambda)] = \beta, \lambda \in Y^*$.

设 $S(y)$ 依 §5.4(1),即

$$S(y) = \inf\{\varphi(x,y) : x \in X\}, y \in Y, \tag{1}$$

称 $S(y)$ 为问题(P)的**值函数**. 由 §5.4(3)有

$$S^*(\lambda) = \varphi^*(0,\lambda), \overline{S}(0) = \beta \leqslant \alpha = S(0). \tag{2}$$

若 φ 凸,则 $S(y)$ 为凸函数. 约定

$$C = \{(y,r) \in Y \times \mathbf{R} : \varphi(x,y) \leqslant r(\exists x \in X)\}. \tag{3}$$

易验证

$$C \subset \text{epi } S \subset \overline{C}. \tag{4}$$

8.3.1 定理 设 α 有限. 假定:(i) 存在 Y 的 0-邻域 V 与 $\sigma > \alpha$,使 $C \cap (V \times (-\infty,\sigma])$ 闭. 则 $S(y)$ 在 $y = 0$ 为 lsc 且问题(P)有解. 条件(i)可代以如下条件之一:(ii) φ 为 lsc,存在 Y 的闭 0-邻域 V 与 $\sigma > \alpha$,使 $\{x \in X : \varphi(x,y) \leqslant \sigma(\exists y \in V)\}$ 为紧集;(iii) X 为自反 B-空间,当 X 中采用弱拓扑时 φ 为 lsc,存在 Y 的闭 0-邻域 V,使得

$$\lim_{r\to\infty} \inf\{\varphi(x,y) : |x| > r, y \in V\} > \alpha. \tag{5}$$

证 若 $S(y)$ 在 $y = 0$ 非 lsc,则有 V° 中的网 $\{y_t\}$ 与 $\varepsilon > 0$,使 $y_t \to 0, S(y_t) < \alpha - \varepsilon$. 于是

$$\begin{aligned}(y_t,\alpha-\varepsilon) &\in (\text{epi } S) \cap (V^\circ \times (-\infty,\alpha)) \\ &\subset \overline{C} \times (V \times (-\infty,\sigma])^\circ \\ &\subset C \times (V \times (-\infty,\sigma])\end{aligned}$$

(用条件(i)),这推出 $(0,\alpha-\varepsilon) \in C$. 由式(3)有 $x \in X$,使 $\varphi(x,0) \leqslant \alpha - \varepsilon$,这矛盾于 $\alpha \leqslant \varphi(x,0)$. 故 $S(y)$ 在 $y = 0$ 为 lsc. 其次,$\forall n \geqslant 1$,取 $x_n \in X$,使 $\varphi(x_n,0) < \alpha + n^{-1}$. 可设 $\alpha + n^{-1} < \sigma$, 于是

$$(0,\alpha + n^{-1}) \in C \cap (V \times (-\infty,\sigma]),$$

这推出 $(0,\alpha) \in C$. 因此有 $\bar{x} \in X$,使 $\varphi(\bar{x},0) \leqslant \alpha$,这表明 $\bar{x}$ 是问题(P)的最优解.

设 V 依条件(ii) 或(iii)，$\sigma > \alpha$ 依条件(ii) 或充分接近于 α、下面证 $C \cap (V \times (-\infty,\sigma])$ 闭. 设 $\{(y_t,r_t)\}$ 是 $C \cap (V \times (-\infty,\sigma]$ 中的网，$(y_t,r_t) \to (y,r)$，则 $(y,r) \in V \times (-\infty,\sigma]$. 只需证 $(y,r) \in C$. 取 $x_t \in X$，使 $\varphi(x_t,y_t) \leqslant r_t \leqslant \sigma$. 若条件(ii) 满足，则不妨设 $x_t \to x$，于是

$$\varphi(x,y) \leqslant \liminf_t \varphi(x_t,y_t) \leqslant r, \tag{6}$$

这推出 $(y,r) \in C$. 若条件(iii) 满足，σ 充分接近于 α，则由式(5) 得出：当 $|x|$ 充分大时对任何 $y \in V$ 有 $\varphi(x,y) > \sigma$. 于是由 $\varphi(x_t,y_t) \leqslant \sigma$ 推出 $\{x_t\}$ 有界. 不妨设 $x_t \rightharpoonup x$，于是同样有(6) 成立，从而 $(y,r) \in C$. 这就证得(ii) $\Rightarrow$ (i) 且(iii) $\Rightarrow$ (i). □

8.3.2 定理 C 为闭集的充要条件是：$S(y)$ 为 lsc，且当 $S(y)$ 有限时 $S(y) = \min \varphi(X,y)$.

证 设 C 闭，则由(4) 知 epi S 闭，因而 $S(y)$ 为 lsc. 若 $S(y)$ 有限，则 $(y,S(y)) \in \text{epi } S = C$，于是有 $x \in X$ 使 $\varphi(x,y) \leqslant S(y)$，从而 $S(y) = \min \varphi(X,y)$.

其次，设 $S(y)$ 为 lsc，且当 $S(y)$ 有限时 $S(y) = \min \varphi(X,y)$. 若 $\{(y_t,r_t)\}$ 是 C 中的网，$(y_t,r_t) \to (y,r)$，则 $S(y_t) \leqslant r_t$，

$$S(y) \leqslant \liminf_t S(y_t) \leqslant r.$$

若 $S(y)$ 有限，则有 $x \in X$，使 $\varphi(x,y) = S(y) \leqslant r$，故 $(y,r) \in C$. 若 $S(y) = -\infty$，则直接由(1) 知有 $x \in X$ 使 $\varphi(x,y) < r$，故亦有 $(y,r) \in C$. 可见 C 为闭集. □

8.3.3 定义 设 $f: X \to [-\infty,\infty]$，$f(x_0)$ 有限，$\varepsilon \in \mathbf{R}$. 称

$$\partial_\varepsilon f(x_0) \stackrel{\text{def}}{=} \{u \in X^* : u \leqslant \Delta f(x_0,\cdot) + \varepsilon\} \tag{7}$$

为 f 在 x_0 的 ε- 次微分. 若 $\forall \varepsilon > 0$：$\partial_\varepsilon S(0) \neq \varnothing$，则说问题(P) 是“近稳定”的(参考 5.4.1).

8.3.4 定理 设 α 有限，则以下条件互相等价：(i) 问题(1) 是近稳定的；(ii) $\alpha = \beta$；(iii) $S(0) = \bar{S}(0)$，$\bar{S}$ 依 5.3.1.

证 (i) $\Rightarrow$ (ii) $\forall \varepsilon > 0$，取 $\lambda \in \partial_\varepsilon S(0)$，则 $\lambda(y) \leqslant S(y) - \alpha + \varepsilon (\forall y \in Y)$. 于是

$$\lambda(y) \leqslant \varphi(x,y) - \alpha + \varepsilon \ (\forall x \in X, y \in Y);$$

$$\begin{aligned} \beta &\geqslant -\varphi^*(0,\lambda) \\ &= -\sup\{\lambda(y) - \varphi(x,y) : x \in X, y \in Y\} \\ &= \inf\{\varphi(x,y) - \lambda(y) : x \in X, y \in Y\} \\ &\geqslant \alpha - \varepsilon. \end{aligned}$$

令 $\varepsilon \downarrow 0$，得 $\alpha = \beta$.

(ii) $\Rightarrow$ (i) 设 $\alpha = \beta$. $\forall \varepsilon > 0$ 取 $\lambda \in Y^*$，使 $-\varphi^*(0,\lambda) < \alpha - \varepsilon$，则对任给 $y \in Y$ 有

$$\begin{aligned} \lambda(y) - S(y) &= \lambda(y) - \inf \varphi(X,y) \\ &= \sup\{\lambda(y) - \varphi(x,y) : x \in X\} \\ &\leqslant \varphi^*(0,\lambda) < -\alpha + \varepsilon, \end{aligned}$$

这推出 $\lambda \in \partial_\varepsilon S(0)$. 可见(P) 是近稳定的.

(ii) $\Leftrightarrow$ (iii) 直接由式(2) 看出. □

8.3.5 定理 设 α 有限，$S(y)$ 凸且在 $y = 0$ 为 lsc，则问题(P) 近稳定.

证 任给 $\varepsilon > 0$，令

$$\begin{aligned} A &= \{(y,r) : S(y) - \alpha + \varepsilon \leqslant r\} \\ &= \operatorname{epi} S + (0, \varepsilon - \alpha). \end{aligned}$$

由 $S(y)$ 凸推出 epi S 与 A 为凸集；由 $S(y)$ 在 $y = 0$ 为 lsc 推出 $(0, 0) \notin \overline{A}$，故有 $(\lambda, \tau) \in Y^* \times \mathbf{R}$，使 $\langle(\lambda,\tau), A\rangle < 0$，即

$$\forall y \in Y : S(y) - \alpha + \varepsilon \leqslant r \Rightarrow \lambda(y) + \tau r < 0. \tag{8}$$

在式(8)中取 $y = 0, r = \varepsilon$ 得 $\tau\varepsilon < 0$. 因此 $\tau < 0$，可设 $\tau = -1$. 若 $S(y) < \infty$，则在不等式 $\lambda(y) < -\tau r = r$ 中令 $r \downarrow S(y) - \alpha + \varepsilon$ 得

$$\lambda(y) \leqslant S(y) - \alpha + \varepsilon.$$

此不等式在 $S(y) = \infty$ 时显然亦成立，因此 $\lambda \in \partial_\varepsilon S(0)$. 这表明问题(P) 是近稳定的. □

注 8.3.5 中的条件"$S(y)$ 凸"可由"φ 凸"推出，而"$S(y)$ 在

$y=0$ 为 lsc”则可代以 8.3.1 中的条件(i) ~ (iii) 之一.

现在将前述结果用到问题

$$\min f(x)=\alpha,\ g(x)\leqslant 0,\ x\in D. \tag{9}$$

令 $L(\cdot,\lambda)=f+\lambda g$,

$$\varphi(x,y)=\begin{cases} f(x), & g(x)+y\leqslant 0, x\in D,\\ \infty, & \text{否则}.\end{cases}$$

在 §5.5 中已指明：问题(9)相当于“$\min\limits_{X}\varphi(x,0)=\alpha$”，而问题(9)的 Rockafellar 对偶即 Lagrange 对偶

$$\max\inf L(D,\lambda)=\beta,\ \lambda\in Y_+^*. \tag{9*}$$

问题(9)的值函数 $S(y)$ 可表成：

$$S(y)=\inf\{f(x):\ g(x)+y\leqslant 0 \text{ 且 } x\in D\}.$$

对 $S(y)$ 可指明以下结论：

1° 若$(f,g)|D$ 类凸或近类凸，则 $S(y)$ 凸或近凸：设 $S(y_i)<r_i(i=1,2)$，取 $x_i\in D$，使 $g(x_i)+y_i\leqslant 0, f(x_i)<r_i(i=1,2)$. 设 $t\in(0,1), x\in D$ 满足

$$f(x)\leqslant tf(x_1)+t'f(x_2), g(x)\leqslant tg(x_1)+t'g(x_2),$$

则 $g(x)+ty_1+t'y_2\leqslant 0$,

$$S(ty_1+t'y_2)\leqslant f(x)<tr_1+t'r_2;$$

令 $r_i\downarrow S(y_i)(i=1,2)$ 得 $S(ty_1+t'y_2)\leqslant tS(y_1)+t'S(y_2)$. 此不等式在 $S(y_1)=\infty$ 或 $S(y_2)=\infty$ 时亦成立.

2° 若 D 闭[弱闭]，$f|D$ 与 $g|D$ 分别为 lsc 与 * lsc[wlsc 与 * wlsc]，则 φ 为 lsc[wlsc].

结合以上结论与 8.3.1 ~ 8.3.5 得出：

8.3.6 推论 设 α 有限. 假定以下条件之一成立：(i) 存在 Y 的 0-邻域 V 与 $\sigma>\alpha$，使

$$\{(y,r)\in Y\times\mathbf{R}:\ y\in V \text{ 且 } f(x)\leqslant r\leqslant\sigma,\ g(x)+y\leqslant 0(\exists x\in D)\}$$

为闭集；(ii) D 闭，$f|D$ 与 $g|D$ 分别为 lsc 与 * lsc，存在 Y 的闭 0-

邻域 V 与 $\sigma > \alpha$,使 $D \cap f^{-1}((-\infty,\sigma]) \cap g^{-1}(V - Y_+)$ 为紧集;
(iii) X 为自反 B-空间,D 弱闭,$f|D$ 与 $g|D$ 分别为 wlsc 与 $*$lsc,$\liminf_{x\to\infty} f(x) > \alpha$,则 $S(y)$ 在 $y = 0$ 为 lsc 且问题(9)有解.

8.3.7 推论 设 α 有限,$(f,g)|D$ 为 $\mathbf{R}_+ \times Y_+$-类凸,且8.3.6的条件(i) ~ (iii)之一满足,则问题(9)是近稳定的,因而 $\alpha = \beta$.

8.3.6 可与 5.2.4,5.2.9 对照;8.3.7 可与 5.5.4(iii)对照.本节结果推广了[115]中的有关结论.

参考文献:[78,85,91,110,113,115].

参考文献

1　Antoni C. Dual problems and separation of sets: Lagrange and Fenchel duality. JOTA, 1997, 93: 227～231

2　Aubin J P. Lipschitz behavior of solutions to convex minimization problems. Math. Oper. Res., 1984, 9: 87～111

3　Aubin J P, Ekeland J. Applied Nonlinear Analysis. Wiley, 1984

4　Auslender A, Cominetti R. A comparative study of multifunction differentiability with applications in mathematical programming. Math. Oper. Res., 1991, 16: 240～258

5　Bector C R, Bector M K. On various duality theorems in nonlinear programming. JOTA, 1987, 53: 509～515

6　Bector C R, Chandra S, Bector M K. Sufficient optimality conditions and duality for quasiconvex programming problems. JOTA, 1988, 59: 209～221

7　Bector C R, Suneja S K, Lalitha C S. Generalized B-vex functions and generalized B-vex programming. JOTA, 1993, 76: 561～576

8　Bender P J. Nonlinear programming in normed spaces. JOTA, 1978, 24: 263～285

9　Ben-tal A. Second-order and related extremality conditions in nonlinear programming. JOTA, 1980, 31: 143～165

10　Ben-tal A, Zowe J. Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems. Math. Program., 1982, 24: 70～91

11　Ben-tal A, Zowe J. A unified theory of first and second order conditions for extremum problems in topological vector spaces. Math. Program. Study, 1982, 19: 39～76

12　Bhakta P C, Roychaudhuri S. Optimization in Banach spaces. J. Math. Anal. Appl., 1988, 134: 460～470

13 Bitran G R. Duality for nonlinear multi-criteria optimization problems. JOTA, 1981, 35: 367～402

14 Bolintineanu S. Necessary conditions for nonlinear suboptimization over the weakly-efficient set. JOTA, 1993, 78: 579～598

15 Borwein J M. Weak tangent cones and optimization in a Banach space. SIAM J. Control Optim., 1978, 16: 512～522

16 Borwein J M. A note on perfect duality and limiting Lagrangeans. Math. Program., 1980, 18: 330～337

17 Borwein J M. Direct theorems in semi-infinite convex programming. Math. Program., 1981, 21: 310～318

18 Borwein J M. Adjoint process duality. Math. Oper. Res., 1983, 8: 403～437

19 Borwein J M, Lewis A. Partially finite convex programming. Math. Program., 1992, 57: 15～48

20 Borwein J M, Strojwas H M. Tangential approximations. Nonl. Anal., 1985, 9: 1347～1366

21 Borwein J M, Strojwas H M. The hypertangent cone. Nonl. Anal., 1989, 13: 125～144

22 Borwein J M, Wolkowicz H. Characterizations of optimality without a constraint qualification for abstract convex program. Math. Program. Study, 1982, 19: 77～100

23 Borwein J M, Wolkowicz H. A simple constraint qualification in infinite dimensional programming. Math. Program., 1986, 35: 83～96

24 Borwein J M, Zhang D. On Fan's minimax theorem. Math. Program., 1986, 34: 232～234

25 Brinkhuis J. Introduction to duality in optimization theory. JOTA, 1996, 91: 523～542

26 Brosowski B. A refinement of an optimality criterion and its applications to parametric programming. JOTA, 1984, 42: 367～382

27 Bu Q-Y, Shen H R. Some properties of efficient solutions for vector optimization. JOTA, 1985, 46: 255～263

28 Burke J V. Second order necessary and sufficient conditions for convex

composite NDO. Math. Program. ,1987,38:287~302

29 Burke J V,Ferris M C,Qian M-J. On the Clarke subdifferential of the distance function of a closed set. J. Math. Anal. Appl. ,1992,166:199~213

30 Burke J V, Ferris M C. Weak sharp minima in mathematical programming. SIAM J. Control Optim. ,1993,31:1340~1359

31 Castellani M, Pappalardo M. First-order cone approximations and necessary optimality conditions. Optimization,1995,35:113~126

32 Chabrillac Y,Crouzeix J-P. Continuity and differentiability properties of monotone real functions of several real variables. Math. Program. Study,1987,30:1~16

33 Chan W L,Huang L R,Ng K F. On generalized second-order derivatives and Taylor expansions in nonsmooth optimization. SIAM J. Control Optim. 1994,32:591~611

34 Chaney R W. Second-order sufficient conditions for nondifferentiable programming problems. SIAM J. Control Optim. ,1982,20:20~33

35 Chaney R W. A general sufficiency theorem for nonsmooth nonlinear programming. Trans. Amer. Math. Soc. ,1983,276:235~246

36 Chaney R W. Second-order necessary conditions in semismooth optimization. Math. Program,1988,40:95~109

37 Chaney R W. Second-order sufficient conditions in nonsmooth optimization. Math. Oper. Res. ,1988,13:660~673

38 陈修素,罗国光. 拓扑向量空间中极值问题的高阶最优性必要条件. 应用数学学报,1990,13:156~167

39 Clarke F H. Optimization and Nonsmooth Analysis. Wiley,1983.

40 Cominetti R, Correa, R. A generalized second-order derivatives in nonsmooth optimization. SIAM J. Control Optim. ,1990,28:789~809

41 Corley H W. Duality theory for maximization with respect cones. J. Math. Anal. Appl. ,1981,84:560~568

42 Corley H W. Existence and Lagrange duality for maximizations of set-valued functions. JOTA,1987,54:489~501

43 Craven B D,Gwinner J,Jeyakumar V. Nonconvex theorems of the

alternative and minimization. Optimization, 1987, 18: 151～163

44 Dancs S. Generalized tangent cone and an optimization problem in a normed space. JOTA, 1990, 67: 43～55

45 Dauer J P, Stadler W. A survey of vector optimization in infinite dimensional spaces, Part 2. JOTA, 1986, 51: 205～241

46 Demyanov V F, Vasilev L V. 不可微最优化. 金洪臻等译. 大连: 大连理工大学出版社, 1991.

47 Dentcheva D, Helbig S. On variational principles, level sets, well-posedness and ε-solutions in vector optimization. JOTA, 1996, 89: 325～349

48 Deumlich R, Elster K H. Φ-conjugation and nonconvex optimization, A survey, Part I. Optimization, 1983, 14: 125～149

49 Deumlich R, Elster K H. Φ-conjugation and nonconvex optimization, A survey, Part Ⅱ. Optimization, 1984, 15: 499～515

50 Deumlich R, Elster K H. Φ-conjugation and nonconvex optimization, A survey, Part Ⅲ. Optimization, 1985, 16: 789～803

51 Dietze S, Schauble M. On the relationship between Fenchel and Lagrange duality for optimization problems in general spaces. Optimization, 1985, 16: 7～14

52 Dolecki S. Abstract study of optimality conditions. J. Math. Anal. Appl., 1980, 73: 24～48

53 董加礼. 不可微多目标优化. 数学进展, 1994, 23: 517～528

54 Dunn J C, Tian T. Variants of the Kuhn～Tucker Sufficient conditions in cones of nonnegative functions. SIAM J. Control Optim., 1992, 30: 1361～1384

55 Egudo R R. Proper efficiency and multiobjective duality in nonlinear programming. J. Inform. Optim. Sci., 1987, 8: 155～166

56 Egudo R R. Efficiency and generalized convex duality for multiobjective programs. J. Math. Anal. Appl., 1989, 138: 84～94

57 Egudo R R, Mond B. Duality with generalized convexity. J. Austral Math. Soc., Ser. B. 1986, 28: 10～21

58 Egudo R R, Hanson M A. Multiobjective duality with invexity. J.

Math. Anal. Appl. ,1987,126:469～477

59　Ekeland I. On the variational principle. J. Math. Anal. Appl. ,1974,47:324～353

60　Elster K H, Thierfelder J. Abstract cone approximations and generalized differentiability in nonsmooth optimization. Optimization, 1988, 19:315～341

61　Facchinei F. Refinements of necessary conditions for Optimality in nonlinear programming. JOTA,1992,73:65～74

62　冯俊文. 非线性规划的Ω共轭对偶理论. 应用数学,1993,6(3):249～255

63　Ferris M C, Mangasarian O L. Minimum principle sufficiency. Math. Program. ,1992,57:1～14

64　Ferro F. An optimization result for set-valued mappings and a stability property in vector problems with constraints. JOTA,1996,90:63～77

65　Fischer Th. On the duality of a nonconvex optimization problem and the strong unicity constant in linear Chebyshev approximation. J. Math. Anal. Appl. ,1992,164:167～177

66　Fletcher R, Watson G A. First and second order condition for a class of nondifferentiable optimization problems. Math. Program. ,1980,18:291～307

67　Floudas C A, Visweswaran V. Primal-relaxed dual global optimization approach. JOTA,1993,78:187～226

68　傅万涛. 关于真有效点集的稠密性. 南昌大学学报,1994,18(3):267～270

69　Fujiwara O. Morse programs: a topological approach to smooth constrained optimization, I. Math. Oper. Res. ,1982,7:602～616

70　Furukawa N, Yoshinaga Y. Higher order variational sets, variational derivatives and higher order necessary conditions in abstract mathematical programming. Bull. Infor. Cybernetics,1988,23:9～40

71　Gajek L, Zagrodny D. Approximate necessary conditions for locally weak Pareto optimality. JOTA,1994,82:49～58

72　Giannessi F. Theorems of the alternative and optimality conditions.

JOTA,1984,42:331～365

73 Giannessi F. Theorems of the alternative for multifunctions with applications to optimization:general results. JOTA,1987,55:233～256

74 Giannessi F. Semidifferentiable functions and necessary optimality conditions. JOTA,1989,60:191～243

75 Glover B M,Jeyakumar V,Oettli W. A Farkas lemma for difference sublinear systems and quasidifferentiable programming. Math. Program.,1994,63:109～125

76 Goberna M,Lopez A. Optimal value function in semi-infinite programming. JOTA,1988,59:261～278

77 Gollan B. Higher order necessary condition for an abstract optimization problem. Math. Program. Study,1981,14:67～76

78 Gollan B. Perturbation theory for abstract optimization problems. JOTA,1981,35:417～441

79 Guerraggio A,Molho E,Zaffaroni A. On the notion of proper efficiency in vector optimization. JOTA,1994,82:1～22

80 Gwinner J. An extension lemma and homogeneous programming. JOTA,1985,47:321～336

81 Gwinner J. Results of Farkas type. Numer. Funct. Anal. Optim.,1987,9:471～520

82 Gwinner J,Oettli W. Theorems of the alternative and duality for inf-sup problems. Math. Oper. Res.,1994,19:238～255

83 Hanson M A. On sufficiency of Kuhn-Tucker condition. J. Math. Anal. Appl.,1981,80:545～550

84 Hanson M A,Mond B. Necessary and sufficient conditions in constrained optimization. Math. Program.,1987,37:51～58

85 Hayashi M,Komiya H. Perfect duality for convexlike programs. JOTA,1982,38:179～189

86 Henrion R. On constraint qualifications. JOTA,1992,72:187～197

87 Hetlich R,Kortanek K O. Semi-infinite programming:theory,methods and applications. SIAM Rev.,1993,35:380～429

88 Hildenbrandt R,Nehse R. On duality-separability relations. Optimiz-

ation,1985,16:805～808

89 Hiriart-Urruty J B. Refinements of necessary optimatlity conditions in nondifferentiable programming, Ⅱ. Math. Program. Study, 1982, 19: 120～139

90 Hiriart-Urruty J B, Seeger A. Calculus rules on a new setvalued second order derivatives for convex functions. Nonl. Anal. ,1989,13:721～728

91 Hiriart-Urruty J B, Phelps R R. Subdifferential calculus using ε-subdifferentials. J. Funct. Anal. ,1993,118:154～166

92 Hoffmann K H, Kornstaedt H J. Higher order necessary conditions in abstract mathematical programming. JOTA,1978,26:533～568

93 胡适耕. 非线性分析. 武汉:华中理工大学出版社,1995.

94 胡适耕. 高阶变分集及其表示. 华中理工大学学报,1996,24(11):1～4

95 胡适耕,沈轶,黄正海. 一个一般的 Motzkin 定理及其应用. 系统科学与数学,1998,18(1):40～46

96 胡晓东. 不可微规划. 运筹学杂志,1988,7(2):23～34

97 胡新生,唐焕文. 一类不可微规划的 Kuhn-Tucker 充分条件. 系统科学与数学,1990,10(2):175～180

98 胡毓达,孙尔江. 多目标规划的几个最优性充分条件. 上海交通大学学报,1994,28(3):89～93

99 黄正海,胡适耕,沈轶. 非光滑多目标规划的广义凸对偶. 华中理工大学学报,1996,24:148～152

100 黄正海,胡适耕,沈轶. 非凸非光滑规划的最优性与对偶性. 华中理工大学学报,1997,25(1):99～102

101 Ito K, Kunisch K. The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces. Math. Program. ,1990,46:341～360

102 Jahn J. Scalarization in vector optimization. Math. Program. ,1984,29: 203～218

103 Jeroslow R G. A limiting Lagrangian for infinitely contrained convex optimization in R^n. JOTA,1981,33:479～495

104 Jeyakumar V. Convex-like alternative theorems and mathematical programming. Optimization,1985,16:643～652

105 Jeyakumar V. A generalization of a minimax theorem of Fan via a theorem of the alternative. JOTA, 1986, 48: 525～533

106 Jeyakumar V. On optimality conditions in nonsmooth inequality constrained minimization. Numer. Funct. Anal. Optim., 1987, 9: 535～546

107 Jeyakumar V. A general Farkas lemma and characterization of optimality for a nonsmooth program involving convex processes. JOTA, 1987, 55: 449～467

108 Jeyakumar V. Equivalence of saddle-points and optima and duality for a class of nonsmooth non-convex problems. J. Math. Anal. Appl., 1988, 130: 334～343

109 Jeyakumar V. Dualtiy and infinite dimensional optimization. Nonl. Anal., 1990, 15: 1111～1122

110 Jeyakumar V. Infinite-dimensional convex programming with application to constrained approximation. JOTA, 1992, 75: 569～586

111 Jeyakumar V. Asymptotic dual conditions characterizing optimality for infinite convex programs. JOTA, 1997, 93: 153～165

112 Jeyakumar V, Glover B M. Nonlinear extensions of Farkas' lemma with applications to global optimization and least squares. Math. Oper. Res., 1995, 20: 818～837

113 Jeyakumar V, Gwinner J. Inequality systems and optimization. J. Math. Anal. Appl., 1991, 159: 51～71

114 Jeyakumar V, Oettli W, Natividad M. A solvability theorem for a class of quasiconvex mappings with applications to optimization. J. Math. Anal. Appl., 1993, 179: 537～546

115 Jeyakumar V, Wolkowicz H. Zero duality gaps in infinite-dimensional programming. JOTA, 1990, 67: 87～108

116 Jeyakumar V, Wolkowicz H. Generalization of Slater's constraint qualification for infinite convex programs. Math. Program., 1992, 57: 85～101

117 Jongen H Th, Twilt F, Weber G W. Semi-infinite optimization: structure and stability of the feasible set. JOTA, 1992, 72: 529～552

118 Karney D F. A dual theorem for semi-infinite convex programs and

their finite subprograms. Math. Program. ,1983,27:75～82

119 Kassay G, Kolumban J. On a generalized sup-inf problem. JOTA, 1996,91:651～670

120 Kaul R N, Kaur S. Optimality criteria in nonlinear programming involving nonconvex functions. J. Math Anal. Appl. ,1985,105:104～112

121 Kaul R N, Suneja S K, Srivastava M K. Optimality criteria and duality in multiple-objective optimization involving generalized invexity. JOTA, 1994,80:465～482

122 Kawasaki H. An envelope-like effect of infinitely many inequ-ality constraints on second order necessary conditions for minimzation problems. Math. Program. ,1988,41:73～96

123 Khanh P Q. Sufficient optimality conditions and duality in vector optimization with invex-convexlike functions. JOTA, 1995, 87: 359～378

124 Khanh P Q, Nuong, T. H. On necessary optimality conditions in vector optimization problems. JOTA, 1988, 58:63～81

125 Kotarski W. Futher generalization of the Dubovickii-Miljutin theorem. JOTA, 1987, 54:565～573

126 Kotarski W. Characterization of Pareto optimal points in problems with multi-equality constraints. Optimization, 1989, 20:93～106

127 Kotarski W. On some specification of the Dubovitskii-Mulyutin theorem for pareto optimal problems. Nonl. Anal. , 1990, 14:287～291

128 Kurcyusz S. On the existence and nonexistence of Lagrange multipliers in Banach spaces. JOTA, 1976, 20:81～110

129 Kyparisis J. Sensitivity analysis for nonlinear programs and variational inequalities with nonunique multipliers. Math. Oper. Res. , 1990, 15: 286～298

130 Ledzewicz-Kowalewska U. On some specification of the Dubovitskii-Milyutin method. Nonl. Anal. , 1986, 10:1367～1371

131 Ledzewicz-Kowalewska U. Application of the method of contractor directions to the Dubovitskii-Milyutin formulism. J. Math. Anal. Appl. ,1987,125:174～184

132 Ledzewicz-Kowalewska U. A necessary condition for the extremal problems under Gâteaux differentiability. J. Math. Anal. Appl. ,1988,134:158～169

133 Ledzewicz-Kowalewska U. Optimality and Pareto optimality conditions for the problems with nonreqular operator equality constraints. Nonl. Anal. ,1991,17:347～360

134 Ledzewicz U,Schaettler H. Second-order conditions for extremum problems with nonregular equality constraints. JOTA,1995,86:113～144

135 李师正.用次微分及法锥表达的对偶问题.应用数学学报,1993,16(3):425～427

136 李师正,王卿文.半无限规划的鞍点与对偶性.数学物理学报,1996,16(2):174～178

137 李仲飞,汪寿阳.多目标规划的 Lagrange 对偶与标量化定理.系统科学与数学,1993,13:211～217

138 李泽民.半无穷向量最优化问题的最优性条件.系统科学与数学,1994,14(4):375～380

139 林锉云,董加礼.多目标优化的方法与理论.长春:吉林教育出版社,1992.

140 Linnemann A. Higher order necessary conditions for infinite and semi-infinite optimizations. JOTA,1982,38:483～511

141 刘三阳.非光滑多目标规划的最优性条件.系统科学与数学,1989,9:53～60

142 刘三阳.非光滑非凸多目标规划的 Wolf 型对偶性.数学研究与评论,1991,11(1):97～101

143 Luc D T. On duality theory in multiobjective programming. JOTA,1984,43:557～582

144 Luc D T. Scalarization of vector optimization problems. JOTA,1987,55:85～102

145 Maeda T. Constraint qualification in multiobjective optimization problems:differential case. JOTA,1994,80:483～500

146 Martein L. Regularity conditions for constrained extremum problems. JOTA,1985,47:217～233

147 Martein L. Lagrange multipliers and generalized differentiable functions in vector extremum problems. JOTA,1989,63:281～297

148 Maruyama Y. Second-order necessary conditions for nonlinear optimization problems in Banach spaces and their application to an optimal control problem. Math. Oper. Res.,1990,15:467～482

149 Maurer H. First- and second-order sufficient optimality conditions in mathematical programming and optimal control. Math. Program. Study,1981,14:163～177

150 Maurer H,Zowe J. First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems. Math. Program.,1979,16:98～110

151 Mclinden L. A minimax theorem. Math. Oper. Res.,1984,9:576～591

152 Mekovsky R R,Ward D E,Upper DSL. Approximates and nonsmooth optimization. Optimization,1990,21:163～177

153 孟志清,邹凯. 集值函数向量优化的 Wolfe 对偶. 湘潭大学学报,1997,19(1)

154 Minami M. Weak Pareto-optimality of multiobjective problems in a locally convex linear topological space. JOTA,1981,34:469～484

155 Minami M. Weak Pareto-optimal necessary conditions in a nondifferentiable multiobjective program on a Banach space. JOTA,1983,41:451～461

156 Mordukhovich B S. Generalized differential calculus for nonsmooth and setvalued mappings. J. Math Anal. Appl.,1994,183:251～287

157 Nemeth A B. A nonconvex vector minimization problem. Nonl. Anal.,1986,10:669～678

158 Noll D. A unified approach to duality in convex programming. J. Math. Anal. Appl.,1991,161:508～521

158 Paeck S. Convexlike and concavelike conditions in alternative,minimax and minimization theorems. JOTA,1992,74:317～332

160 Palato J. A survey of conical approximations used in the optimization. Optimization,1989,20:147～162

161 Papageorgiou N S. Nonsmooth analysis on partially ordered vector

spaces: the subdifferential theory. Nonl. Anal. ,1986,10:615～637

162 Pappalardo M. Tangent cones and Dini derivatives. JOTA,1991,70:97～107

163 Pardalos P M,Rosen J B. Methods for global concave minimization: a bibliographic survey. SIAM Rev. ,1986,28:367～379

164 Pascoletti A, Serafini P. Scalarizing vector optimization problems. JOTA, 1984,42:499～524

165 Pellegrini. On a Lagrangian sufficient optimality condition. JOTA, 1991,68:19～33

166 Penot J P. A characterization of Clarke's strict tangent cone via non linear semigroups. Proc. Amer. Math. Soc. 1985,93:128～132

167 Polak E. On the mathematical foundations of nondifferentiable optimization in engineering design. SIAM Rev. ,1987,29:23～89

168 Pomerol J Ch. The Lagrange multiplier set and the generalized gradient set of the marginal function of a differentiable program in a Banach space. JOTA,1982,38:307～317

169 Pomerol J Ch. Inequality systems and minimax theorems. J. Math. Anal. Appl. ,1984,103:263～292

170 Preda V. On efficiency and duality for multiobjective programs. J. Math. Anal. Appl. ,1992,166:365～377

171 Preda V. On sufficiency and duality for generalized quasiconvex programs. J. Math. Anal. Appl. ,1994,181:77～88

172 Quan X C. A two-function minimax theorem. J. Math Anal. Appl. , 1994,185:660～674

173 Reiland T W. Optimality conditions and duality in contionuous programming I: convex programming and a theorem of the alternative. J. Math. Anal. Appl. ,1980,77:297～325

174 Reiland T W. A geometric approach to nonsmooth optimization with sample application. Nonl. Anal. ,1987,10:1169～1184

175 Revalski J P,Zhivkov N V. Well-posed constrained optimization problems in metric spaces. JOTA,1993,76:145～163

176 Robinson S M. Local structure of feasible sets in nonlinear progr-

amming. Part Ⅱ: nondegeneracy. Math. Program. Study, 1984, 22: 217 ～230

177 Rockafellar R T. Convex Analysis. 2ed. Univ. Press, 1972.

178 Rockafellar R T. Generalized directional derivatives and subgradients of nonconvex functions. Canada J. Math., 1980, 32: 257～280

179 Rockafellar R T. Extensions of subgradient calculus with applications to optimization. Nonl. Anal., 1985, 9: 665～698

180 Rockafellar R T. First and second-order epi-differentiability in nonlinear programming. Trans. Amer. Math. Soc., 1988, 307: 75～108

181 Rockafellar R T. Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives. Math. Oper. Res., 1989, 14: 462～484

182 Rodrigues B, Simons S. Conjugate functions and subdifferentials in non-normed situations for operators with complete graphs. Nonl. Anal., 1988, 12: 1069～1078

183 Rueda N, Hanson M A. Optimality criteria in mathematical programming involving generalized invexity. J. Math. Anal. Appl., 1988, 130: 375～385

184 Rueda N, Hanson M A, Singh C. Optimaty and duality with generalized convexity. JOTA, 1995, 86: 491～500

185 Schecter S. Structure of the first-order solution set for a class of nonlinear programs with parameters. Math. Program., 1986, 34: 84～110

186 Schirotzek W. On a theorem of Ky Fan and its application to nondifferentiable optimization. Optimization, 1985, 16: 353～366

187 Schirotzek W. Nonasymptotic necessary conditions for nonsmooth infinite optimization problems. J. Math. Anal. Appl., 1986, 118: 535～546

188 Shapiro A, Al-Khayyal F. First-order conditions for isolated locally optimal solutions. JOTA, 1993, 77: 189～196

189 史树中. 非光滑分析. 数学进展, 1986, 15(1): 9～21

190 施保昌, 周晓阳. 非线性规划的不使用导数的解析方法. 应用数学,

1992,5(4):18~24

191 沈铁,胡适耕,黄正海.广义凸集与广义凸函数.华中理工大学学报,1997,25:86~89

192 沈铁,王书宁,胡适耕,廖晓昕.鞍点存在定理与 minimax 定理.华中理工大学学报,1997,25:37~39

193 Shimizu K,Aiyoshi E,Katayama R. Generalized Farkas' theorem and optimization of infinitely constrained problems. JOTA,1983,40:451~462

194 Shioji N,Takahashi W. Fan's theorem concerning systems of convex inequalities and its applications. J. Math. Anal. Appl.,1988,135:383~393

195 Singer I. A general theory of dual optimization problems. J. Math. Anal. Appl.,1986,116:77~130

196 Singer I. Some general Lagrangian duality theorems. J. Math. Anal. Appl.,1989,144:26~51

197 Singer I. Some further duality theorems for optimization problems with reverse convex constraint sets. J. Math Anal. Appl.,1992,171:205~219

198 Singh C. Optimality conditions in multiobjective differentiable programming. JOTA,1987,53:115~123

199 Singh C,Bhatia D,Rueda N. Duality in nonlinear multiobjective programming using augmented Lagrangian functions. JOTA,1996,88:659~670

200 Sion M. On general minimax theorems. Pacif. J. Math.,1958,8:171~176

201 Song W. Lagrangian duality for minimization of nonconvex multifunctions. JOTA,1977,93:167~182

202 Stefanescu A. A general min-max theorem. Optimization,1985,16:497~504

203 Still G,Streng M. Optimality conditions in smooth nonlinear programming. JOTA,1996,90:483~515

204 Studniarski M. Mean value theorems and sufficient optimality con-

ditions for nonsmooth functions. J. Math. Anal. Appl. ,1985,111:313 ~326

205 Swartz C. A general Farkas lemma. JOTA,1985,46:237~244

206 Tamminen E V. Sufficient conditions for the existence of multipliers and Lagrangian duality in abstract optimization problems. JOTA,1994, 82:93~104

207 Tanaka T. Existence theorems for cone saddle point of vectorvalued functions in infinite dimensional spaces. JOTA,1989,62:127~138

208 Tanaka T. Generalized quasiconvexities,cone saddle points and minimax theorem for vector-valued functions. JOTA,1994,81:355~377

209 Tanino T. Sensitivity analysis in multiobjective optimization. JOTA, 1988,56:479~499

210 Tanino T,Sawaragi Y. Duality theory in multiobjective programming. JOTA,1979,27:509~529

211 Tanino T,Sawaragi Y. Conjugate maps and duality in multiobjective optimization. JOTA,1980,31:473~479

212 Thach P T. Convex minimization under Lipschitz constraints. JOTA, 1990,64:595~614

213 Thibault L. On generalized differentials and subdifferentials of Lipschitz vector-valued functions. Nonl. Anal. 1982,6:1037~1053

214 Thibault L. Subdifferentials of nonconvex vector-valued functions. J. Math. Anal. Appl. ,1982,86:319~344

215 Thibault L. Tangent cones and quasi interior tangent cones to multifunctions. Trans. Amer. Math. Soc. ,1983,277:601~621

216 Tind J,Wolsey L. A. An elementary survey of general duality theory in mathematical programming. Math. Program. ,1981,21:241~261

217 Treiman J C. Characterization of Clarke's tangent and normal cones in finite and infinite dimension. Nonl. Anal. ,1983,7:771~783

218 Treiman J C. Shrinking generalized gradients. Nonl. Anal. ,1988,12: 1429~1450

219 Treiman J C. An infinite class of convex tangent cones. JOTA,1991, 68:563~581

220 Tuy H. Convex programs with an additional reverse convex constraint. JOTA,1987,52:463～472

221 Ursescu C. Tangent set's calculus and necessary conditions for extremality. SIAM J. Control Optim.,1982,20:563～574

222 Vial J P. Strong and weak convexity of sets and functions. Math. Oper. Res.,1983,8:231～259

223 Walczek S. Some properties of cones in normed spaces and their applications to investigating extremal problems. JOTA,1984,42:561～582

224 Wang S,Li Z. Scalarization and Lagrange duality in multiobjective optimization. Optimization,1992,26:315～324

225 Warburton A R. Quasiconcave vector maximization: connectedness of the set of Pareto-optimal and weak Pareto-optimal alternatives. JOTA, 1983,40:537～557

226 Ward D E. Convex subcones of the contingent cone in nonsmooth calculus and optimization. Trans. Amer. Math. Soc.,1987,302:661～682

227 Ward D E. Isotone tangent cones and nonsmooth optimization. Optimization,1987,18:769～783

228 Ward D E. Chain rules for nonsmooth functions. J. Math. Anal. Appl., 1991,158:519～538

229 Ward D E, Borwein, J M. Nonsmooth calculus in finite dimensions. SIAM J. Control Optim.,1987,25:1312～1340

230 Watkins G G. Nonsmooth Milyutin-Dubovitskii theory and Clarke's tangent cone. Math. Oper. Res.,1986,11:70～80

231 Weinstein S E, Xu Yuesheng. A duality approach to best uniform convex approximation. J. Math. Anal. Appl.,1991,160:314～322

232 Weir T, Mond B. Generalized convexity and duality in multiobjective programming. Bull. Amer. Math. Soc.,1989,39:287～299

233 Xu Zengkun. Generalization of nonhomogeneous Farkas lemma and applications. J. Math. Anal. Appl.,1994,186:726～734

234 Xu Zengkun. Mixed type duality in multiobjective programming. J. Math. Anal. Appl.,1996,198:621～635

235 Ye Y L. D-invexity and optimality conditions. J. Math. Anal. Appl.,

1991,162:242～249

236 Yosida K. Functional Analysis. Springer-Verlag,1978

237 Yuan Yaxiang. Nonlinear programming ～ a review. Advances in Math.,1990,19:12～24

238 Zeidler E. Nonlinear Functional Analysis and its Applications, Ⅲ. Springer-Verlag,1986

239 Zhang J,Mond B. Duality for a nondifferentiable problem. Bull. Austral Math. Soc.,1997,55:29～44

240 张石生,舒永录.多值映射的变分不等式及其对非线性规划和鞍点问题的应用.应用数学学报,1991,14(1):32～39

241 郑喜印,Arrow-Barankin-Blackwell 定理在局部凸拓扑向量空间中的一般形式.数学年刊,1997,18A(5):659～664

242 Zowe J, Kurcyusz S. Regularity and stability for the mathematical programming problems in Banach spaces. Appl. Math. Optim.,1979,5:49～62

名词索引

名词后的数字表示该名词最初出现的页数；当标出几个页数时，表明该名词在几个不同的意义上使用.

1～3 画

8 画

9 画

10 画

11 画

12 画及以上